

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМ. ІГОРЯ СІКОРСЬКОГО”
Факультет електроніки
Кафедра електронної інженерії

До захисту допущено
Завідувач кафедри
В. І. Тимофєєв
“ ” червня 2020 р.

Дипломна робота
освітнього рівня «бакалавр»
за спеціальністю 153 мікро- та наносистемна техніка

на тему: «Використання штучних нейронних мереж для діагностики
хвороби Альцгеймера на основі зображень МРТ»

Виконав студент 4 курсу, групи ДМ-61

Палій Олександр Олександрович

(прізвище, ім'я, по батькові)

(підпис)

Керівник доц. каф. ЕІ, доц., к.т.н. А.О. Попов

(посада, вчене звання, науковий ступінь, прізвище та ініціали)

(підпис)

Рецензент доц. каф. ЕПС, доц., к.т.н. К. С. Клен

(посада, вчене звання, науковий ступінь, прізвище та ініціали)

(підпис)

Засвідчую, що у цій дипломній роботі немає
запозичень з праць інших авторів без відповідних
посилань.

Студент

(підпис)

Київ – 2020 року

**Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”**

Факультет електроніки
Кафедра електронної інженерії
Освітній рівень «бакалавр»
за спеціальністю 153 мікро- та наносистемна техніка

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ В. І. Тимофєєв
“__” _____ червня 2020 р.

**З А В Д А Н Н Я
НА ДИПЛОМНУ РОБОТУ СТУДЕНТУ**

Палію Олександр Олександровичу
(прізвище, ім'я, по батькові)

1. Тема роботи: «Використання штучних нейронних мереж для діагностики хвороби Альцгеймера на основі зображень МРТ»

керівник роботи доц. ЕІ, доц., к.т.н. А.О. Попов _____,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від “ 25 ” травня 2020 року № 1196-с

2. Строк подання студентом роботи: 12.02.2020 р.

3. Вихідні дані до роботи: база даних знімків магнітно-резонансної томографії головного мозку людини

4. Зміст дипломної роботи (перелік питань, які потрібно розробити): застосування методів машинного навчання в аналізі знімків МРТ, застосування методів глибокого навчання в аналізі знімків МРТ, огляд основних концепцій в діагностиці нейродегенеративних захворювань.

5. Перелік графічного (ілюстративного) матеріалу (із зазначенням обов'язкових креслень, плакатів, презентацій тощо): рисунки в тексті пояснювальної записки, ілюстрації до захисту.

6. Дата видачі завдання: 09.04.2020 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів виконання дипломної роботи	Строк виконання етапів роботи	Примітка
1	Формування бази даних	09.04.20 – 09.05.20	
2	Виділення шуму на знімках МРТ і його фільтрація, приведення зображень до необхідного формату	10.05.20 – 26.05.20	
3	Проектування штучної нейронної мережі	27.05.20 – 10.06.20	
4	Навчання штучної нейронної мережі	11.06.20 – 15.06.20	
5	Проведення серії тестів навченої моделі штучної нейронної мережі на тестовій підвибірці з бази даних знімків МРТ	16.06.20 – 31.06.20	
6	Підготовка до захисту	01.07.20 – 12.02.20	

Студент



(підпис)

Палій О. О.

(прізвище та ініціали)

Керівник роботи



(підпис)

Попов А. О.

(прізвище та ініціали)

РЕФЕРАТ

Дипломна робота: 80 с., 5 ч., 6 табл., 35 рис., 57 джерел.

НЕЙРОННІ МЕРЕЖІ, ХВОРОБА АЛЬЦГЕЙМЕРА, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, RESNET50V2

Об'єктом розгляду є діагностика хвороби Альцгеймера. Предмет роботи – застосування алгоритмів машинного навчання в класифікації зображень, використання штучний нейронних мереж в діагностиці хвороби Альцгеймера.

Метою роботи є розробка на основі штучних нейронних мереж системи автоматизованої діагностики хвороби Альцгеймера на основі знімків МРТ.

Відсутність ефективних методів діагностики і лікування нейродегенеративних захворювань, поряд з постійним збільшенням їх кількості у світовій популяції, призводить щорічно до величезних економічних втрат. Методи машинного навчання, які демонструють величезні успіхи в різних областях, є одними з головних кандидатів для революції в сфері діагностики.

В роботі, послідовно, проведено розгляд основних підходів діагностики, основного провокуючого фактору розвитку деменції - хвороби Альцгеймера, з акцентом на використання знімків МРТ головного мозку людини і методів глибокого навчання для їх аналізу. Зроблено висновок про відсутність знань першопричини цієї хвороби, як головної складності її діагностики. Окремо, викладені основні концепції машинного навчання та глибокого навчання, зокрема.

Була розроблена система автоматичної діагностики хвороби Альцгеймера на основі знімків МРТ для виділення трьох класів пацієнтів: здорові, з легкої деменцією через хворобу Альцгеймера, з хворобою Альцгеймера. Система була побудована на основі нейромережевої моделі ResNet50V2.

ABSTRACT

BSc thesis: 80 pages, 5 hour, 6 tables, 35 figures, 57 sources.

NEURAL NETWORKS, ALZHEIMER'S DISEASE, ROLLED NEURAL NETWORKS, RESNET50V2

The diagnosis is based on the diagnosis of Alzheimer's disease. The subject of work - the current state of diagnosis of Alzheimer's disease, the use of machine learning algorithms in the classification of images, the use of artificial neural networks in the diagnosis of Alzheimer's disease.

The aim of the work is to develop on the basis of artificial neural networks a system of automated diagnosis of Alzheimer's disease based on MRI images.

Deep learning reached unprecedented heights in tasks that were considered subject only to man. Not so long ago it was impossible to imagine that a computer could surpass a person in the accuracy of image classification. But this happened in 2014, when the neural network won, then a student at Stanford University, Andrei Karpaty, in a competition to recognize 200 dog breeds. But the triumphal procession of neural networks did not end there. And already in 2017, the neural network defeated the man again, this time she won two games, out of the planned three, the strongest player in Go - Ke Jie. And these are just a few areas of activity where a person cedes his superiority to neural networks. Now the scope of neural networks is unusually wide: from the classification of images to predicting the rate of melting of ice masses in Antarctica, from the generation of automatic responses on the Internet to predicting the evolution of galaxies. Therefore, it is not surprising that now technologies based on neural networks are closely integrated into human everyday life. Sometimes the person himself is not aware of this.

But although neural networks have already reached unprecedented heights, their application in the healthcare sector is still extremely insignificant. This happens due to legal restrictions, a biased attitude towards systems whose internal content is not known, and because of inertia, underfunding in some countries of the healthcare system, and in some cases, the lack of convincing results of neural networks in many areas. But even the

last circumstance is not critical for many areas, since in principle there are no other effective methods. In particular, such an area is the diagnosis of Alzheimer's disease, where existing methods can determine its presence only in the late stages, when existing treatment strategies turn out to be powerless. Therefore, an attempt to introduce a method that will allow you to gain time in the disease for the patient in the conditions, as mentioned above, of the absence of other alternatives, is, as they say, a win-win option.

Therefore, in the thesis, a comprehensive review of the diagnosis of Alzheimer's disease was carried out, with an emphasis on the use of neural networks in the diagnosis of Alzheimer's. The quintessence of work was the creation, on the basis of preliminary studies both in the field of neural networks and the use of neural networks in the diagnosis of Alzheimer's disease, an automated system for diagnosing Alzheimer's disease. For a clear and comprehensive coverage of the state of diagnosis of Alzheimer's disease and the existing capabilities of piece neural networks, the structure of the thesis was chosen, which consists of seven sections. In each of them, all aspects of the current state of diagnosis of Alzheimer's disease, the concept of neural networks, in particular, and machine learning, in general, were consistently presented.

In the first section, entitled "Introduction", the justification of the growing social importance of the problem of Alzheimer's was carried out through the prism of economic litigation, which every year is carried by the global economies and budgets of specific families in which there are people with Alzheimer's disease. Objective indicators of the quality of life faced by people with Alzheimer's disease were given. The forecast was made on the basis of open studies of the growth of the economic burden on the world economies if the current situation in the sphere does not change and Alzheimer's disease is not diagnosed in the early stages or effective drugs are not invented to combat it. In general, in the introductory part, the need is formulated to create an automated system for the early diagnosis of Alzheimer's disease.

The second part of the thesis was an introduction to the basic concepts of neurodegenerative diseases and Alzheimer's disease, in particular. The difference between dementia and Alzheimer's disease has been carefully described, due to the fundamental difference between these concepts and the fact that in an environment remote

from the medical, these concepts can often be identified or confused with each other, which in the end result can lead to incorrect interpretation of some research results. and possibly misinterpretation of scientific results. What can cause unintentional falsification of scientific results.

An existing algorithm for diagnosing Alzheimer's disease was also described, where it was noted that there are no reliable methods that would determine exactly Alzheimer's disease. But there are only a number of methods that, with varying degrees of reliability, should, as they say, give indirect evidence, in this case, for other diseases, to establish Alzheimer's disease. The emphasis is placed on the technological backwardness of the methods for diagnosing Alzheimer's disease, in comparison with modern means available to medicine. The fact of insufficient efficiency of using modern technical means, such as magnetic resonance imaging and positron emission tomography, was also noted. The stages of development of Alzheimer's disease are described.

At the end of the section, a conclusion was drawn, in which, based on the section, the main difficulties associated with the diagnosis of Alzheimer's were summarized. Thus, the introduction makes a complete introduction to the basic concepts and formulates the necessary terminological basis for the diagnosis of Alzheimer's disease.

A section called "Machine Learning" is intended to give a soft introduction to machine learning, to introduce its basic concepts.

It describes the history of the creation of machine learning algorithms from first ideas to algorithms running on supercomputers. In parallel with this, the main tasks are introduced, which should be solved by machine learning algorithms. A formal statement of the problem of machine learning. A soft introduction is made that machine learning is nothing more than the task of approximating the objective function, with the problems inherent in the problem of approximating the objective function. Accordingly, the analysis of methods for approximating the objective functions is carried out. In the course of which the problem of retraining machine learning systems and methods of dealing with this phenomenon are described.

The section "Piece neural networks", as well as the section "Machine learning", describes the history of the creation of the science of piece neural networks, at the same

time introduces and describes the basic concepts of piece neural networks, such as: piece neuron, activation function, and others. A comparative analysis of a biological neuron and a piece neuron is carried out. The current state of the field of deep learning is described in detail, various kinds of information are provided on the scope of the use of piece neural networks, on the competition of piece neural networks with humans, on the market for the use of piece neural networks and its growth. A full description of the first algorithms built on piece neural networks is made, in particular, a detailed description of one of the first neural network algorithms, the Rosenblatt perceptron, is carried out.

The two most used architectures for image analysis are introduced, namely, a fully connected piecewise neural network and a convolutional piecewise neural network. After this, a complete derivation of the most commonly used method in the training of neural networks is done - the method of back propagation of errors.

In the section "Use of neural networks in the diagnosis of Alzheimer's disease" a detailed review of the currently existing methods is carried out, with a description of the procedure for working with them, the use of neural networks in the diagnosis of Alzheimer's disease. The description of all the methods of using piece neural networks in the diagnosis of Alzheimer's disease provides links to scientific articles that reveal the essence of each of them at the level of the experiment. The emphasis is on the advantages and disadvantages of using convolutional architectures such as: VGGNet, ResNet, ResNet50V2 and so on. The main obstacles for creating an effective diagnostic system based on piece neural networks are formulated.

In the next section, entitled "Description of the developed scheme", in the first part, the rationale and a detailed description of the selected model for creating an automated diagnosis of Alzheimer's disease are made. The second part describes in detail, based on the experiments, the functionality of the developed system for the automatic diagnosis of Alzheimer's disease, namely, to evaluate the accuracy and stability of the system for the automatic diagnosis of Alzheimer's disease.

The final part of the thesis entitled "Conclusions" summarizes the work done, gives the main functional characteristics of the developed system for the automatic diagnosis of Alzheimer's disease, compares the developed system for the automatic diagnosis of

Alzheimer's disease with other automatic systems for the diagnosis of Alzheimer's disease, describes the main ways to improve the developed system for the automatic diagnosis of Alzheimer's disease and a conclusion is drawn on the practical value of the developed system for diagnosing Alzheimer's disease.

ЗМІСТ

ВСТУП.....	12
1 ОСНОВНІ ПОНЯТТЯ НЕЙРОДЕГЕНЕРАТИВНИХ ХВОРОБ	14
1.1 Деменція та хвороба Альцгеймера.....	14
1.2 Стадії хвороби Альцгеймера	15
1.3 Основні методи діагностики хвороби Альцгеймера	18
1.4 Висновок до першого розділу	20
2 МАШИННЕ НАВЧАННЯ	22
2.1 Історичні передумови виникнення систем машинного навчання	22
2.2 Постановка задачі машинного навчання	24
2.3 Методи оптимізації	25
2.4 Проблеми машинного навчання.....	28
2.5 Алгоритми машинного навчання	28
2.6 Стратегії машинного навчання	29
2.7 Висновок до другого розділу	30
3 ШТУЧНІ НЕЙРОННІ МЕРЕЖІ	31
3.1 Сучасний стан методів глибокого навчання	31
3.2 Основні поняття нейронних мереж та історія їх виникнення	33
3.3 Порівняння біологічного і штучного нейрона	36
3.4 Перші нейронні алгоритми (перцептрон Розенблатта).....	37
3.5 Метод зворотного поширення похибки.....	40
3.6 Повнозв'язані нейронні мережі.....	45
3.8 Висновок до третього розділу	47
4 ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ В ДІАГНОСТИЦІ	48
ХВОРОБИ АЛЬЦГЕЙМЕРА.....	48
4.1 Основні стратегії проектування систем діагностики	48
4.3 Проблеми в автоматичній класифікації знімків МРТ	51
4.4 Висновок до четвертого розділу	52

5	ОПИС РОЗРОБЛЕНОЇ СИСТЕМИ.....	53
5.1	Опис архітектури	53
5.3	Висновки до п'ятого розділу	73
	ВИСНОВКИ	75
	ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	76

ВСТУП

Хвороба Альцгеймера (ХА) - нейродегенеративна хвороба, яка вважається однією з головних проблем сфери охорони здоров'я у майбутньому, зараз на неї припадає від 60% до 80% усіх випадків деменції [1]. Вона є шостою провідною причиною смертності у США. За оцінками асоціації Альцгеймера зараз неї страждає близько 5,7 мільйонів американців, а до 2050 року їх кількість може збільшитися до 14 мільйонів [2]. За прогнозами Всесвітньої організації охорони здоров'я (ВОЗ), частка людей в світі з хворобою Альцгеймера старше 65 років повинна збільшитися із 6,8% у 2005 році до 16,2% у 2050 році [3]. Очікується, що до 2040 року смертність від хвороби складе 81млн осіб в світі, при відсутності змін в рівні захворюваності, смертності, тобто без появи нових лікувальних та профілактичних засобів [4].

Різноманітні епідеміологічні дослідження показали, що імовірність виявлення хвороби має експоненціальну залежність від віку пацієнта. Так в віці 65-69 років розповсюдженість хвороби серед населення складає 1,5-2,5%, досягаючи показника 40% в віці 90-94 років [5, 6]. Медіана тривалості життя для пацієнтів з моменту встановлення діагнозу у віці 65 років становить 9 років, у віці 90 років – 3 роки [7].

До того ж, наростаюча криза хвороби Альцгеймера несе великі економічні втрати як для країн, так і для бюджетів сімей в яких є хворі. Ці втрати поділяються на прямі (забезпечення лікарняними ресурсами, медичними послугами, соціальними послугами) і опосередковані (втрата прибутку пацієнтом, втрата або зниження прибутку його піклувальниками, через втрату робочого часу чи негативний психо-емоційний стан).

Так після онкологічних захворювань і ішемічної хвороби серця, ХА є третім найдорожчим захворюванням в США [8]. Загальні світові витрати, пов'язані з боротьбою та профілактикою хвороби Альцгеймера в 2005 році склали 315млн

доларі, причому 70% цих витрат припадало на розвинені країни [9]. Затрати на деменцію складають 66%-75% від бюджету сім'ї [10].

Хоча зараз немає терапії, яка діє саме на причину хвороби Альцгеймера, тим самим зупиняючи її прогресування, все ж існують препарати, які здатні покращити загальний стан пацієнта, діючи на інші фактори старіння [11]. Застосування їх на ранніх стадіях може відтермінувати момент повної втрати дієздатності, тим самим зменшити затрати сімейного бюджету. Але зараз виявлення хвороби на ранній стадії є майже не можливим, через відсутність ефективних методів її діагностики. Втрата моменту початку хвороби в подальшому може стати причиною помилок на виробництві, різних юридичних проблем, коли буде констатована втрата дієздатності.

Враховуючи вище наведене, можна констатувати, що розробка системи автоматичної діагностики хвороби Альцгеймера, яка і є метою цієї роботи, має велику соціальну значимість та здатна зменшити щорічні витрати на лікування та профілактику хвороби Альцгеймера.

1 ОСНОВНІ ПОНЯТТЯ НЕЙРОДЕГЕНЕРАТИВНИХ ХВОРОБ

1.1 Деменція та хвороба Альцгеймера

Деменція являє собою найбільш загальне визначення для всіх станів, причиною яких є деградація нервової тканини, і які проявляються в погіршенні як вищих психічних функцій, таких як пізнання, пам'ять, так і функцій, залежних від діяльності нервової системи, зокрема, координація рухів.

Хвороба Альцгеймера є тільки одним варіантом з багатьох, які можуть призводити до деменції, хоч на неї і припадає близько 60-80% всіх випадків виникнення деменції [1].

Сам же процес постановки діагнозу хвороби Альцгеймера є комплексною процедурою. З однієї сторони, такий діагноз виставляється на підставі скарг пацієнта і проходження ним психіатричного інтерв'ю, за якими і робиться висновок про прогресуюче погіршення вищих психічних функцій. Тобто умовою для постановки діагнозу, є наявність у пацієнта прогресуючого погіршення вищих психічних функцій. Тут, і це особливо важливо, наявність прогресуючого стану є необхідною умовою. Але проходження інтерв'ю і збір анамнезу в знайомих дозволяє отримати тільки інформацію про симптоми, а не про причини. Саме тому після нього не можливо встановити діагноз. Так як прогресуюче погіршення вищих психічних функцій може бути викликане й іншими захворюваннями, зокрема, судинною деменцією. Її першопричиною є порушення кровопостачання нервової тканини, яке виникає в результаті різного роду інсультів, формування склеротичних бляшок. Але що більш важливо для діагностики, ці причини легко визначаються, наприклад, на знімках МРТ. Сама ж причина виникнення хвороби Альцгеймера до кінця не визначена, і навіть немає супутніх їй артефактів, які мали б сто відсоткову кореляцію з нею. Хоча є деякі артефакти, зокрема, амілоїдні бляшки та нейрофібрилярні клубки, які доволі часто можуть зустрічатися при ній [12, 13]. Доцільно сказати, що зараз діагноз хвороби Альцгеймера виставляється в більшій мірі на основі виключення у пацієнта всіх можливих станів, які

проявляться симптомами характерними для хвороби Альцгеймера, а також мають чітко встановлені причини, наприклад, судинна деменція. Тобто, якщо є прогресуюче погіршення вищих психічних функцій, а причина не встановлена, тоді є доцільність говорити про наявність хвороби Альцгеймера. Також із-за невизначеності причин, зараз не можна стверджувати, що ХА має лише одну причину виникнення. На користь багатопричинної теорії може свідчити відмінна реакція, від сильної до резистивної, на існуючі препарати серед різних людей з підтвердженим діагнозом [14]. Тому цілком можливо, що ХА повторить історію, наприклад, цукрового діабету, який після проведення додаткових досліджень став представляти групу окремих захворювань із схожою симптоматикою, зокрема, діабет 1-шого типу, діабет 2-гого типу і так далі.

1.2 Стадії хвороби Альцгеймера

Зараз виділяють п'ять стадій розвитку хвороби Альцгеймера [15]:

- доклінічна хвороба Альцгеймера;
- легке когнітивне порушення через хворобу Альцгеймера;
- легка деменція через хворобу Альцгеймера;
- помірна деменція через хворобу Альцгеймера;
- важка деменція через хворобу Альцгеймера;

Доклінічна хвороба Альцгеймера.

Хвороба Альцгеймера починається задовго до появи будь-яких симптомів. Зазвичай вона визначається тільки в умовах досліджень. На цій стадії ні ви, ні оточуючі не помічають симптомів.

Легке когнітивне порушення (ЛКП) через хворобу Альцгеймера.

Люди з ЛКП мають невеликі зміни в своїй пам'яті і мисленні. Ці зміни не настільки значні, щоб вплинути на роботу або відносини. Люди з ЛКП можуть мати провали в пам'яті, коли справа доходить до інформації, яку зазвичай легко запам'ятати, таку як розмови, недавні події чи зустрічі.

Легка деменція через хворобу Альцгеймера.

Хвороба Альцгеймера часто діагностується на стадії слабкої деменції, коли сім'ї і лікарям стає зрозуміло, що у людини виникають серйозні проблеми з пам'яттю і мисленням, що впливають на повсякденну діяльність.

На стадії легкої деменції люди можуть відчувати:

- **Втрата пам'яті про останні події.** Людям може бути особливо важко запам'ятовувати тільки що отриману інформацію, через що люди можуть задавати одне і те саме питання знову і знову.
- **Труднощі з рішенням проблем, складними завданнями і здоровими судженнями.** Багато людей відчують помилки в судженнях, наприклад, при прийнятті фінансових рішень.
- **Зміни в особистості.** Люди можуть стати пригніченими або замкнутими, особливо в соціально складних ситуаціях, або проявляти нехарактерну дратівливість або гнів. Зниження мотивації для виконання завдань також є поширеним явищем.
- **Складність в організації та вираженні думок.** Пошук відповідних слів для опису об'єктів або чіткого вираження ідей стає все більш складним завданням.
- **Заблукати або втратити речі.** Людям все важче знаходити дорогу, навіть у знайомих місцях. Також часто втрачають речі.

Помірна деменція через хворобу Альцгеймера.

Під час стадії помірної деменції при хворобі Альцгеймера люди стають більш розгубленими і забудькуватими і починають потребувати більшої допомоги в повсякденних справах і турботі про себе.

Люди з помірною стадією деменції при хворобі Альцгеймера можуть:

- **Погіршення судження і поглиблення плутанини.** Люди втрачають зв'язок з місцем і часом в якому вони знаходяться. Вони можуть сплутати членів сім'ї або близьких друзів один з одним або прийняти незнайомих за сім'ю, можуть бродити, можливо, в пошуках обстановки, яка здається більш знайомою. Ці труднощі збільшують небезпеку знаходження цих людей в самотності.
- **Зазнають ще більшої втрати пам'яті.** Люди можуть забути подробиці своєї особистої історії, наприклад, адресу або номер телефону, місце відвідування школи. Вони повторюють улюблені історії або складають історії, щоб заповнити прогалини в пам'яті.
- **Потрібна допомога з деякими повсякденними справами.** Може знадобитися допомога у виборі відповідного одягу для даної ситуації або погоди, а також при купанні, догляді за собою і інших способах догляду за собою. Деякі люди іноді втрачають контроль над сечовим міхуром або випорожненнями.
- **Проходять значні зміни в особистості і поведінці.** Для людей з помірною стадією деменції немає нічого незвичайного в тому, щоб розвинути необґрунтовані підозри, наприклад, переконатися в тому, що у них крадуть друзі, родичі або професійні піклувальники.

Важка деменція через хворобу Альцгеймера

На цій стадії психічна функція продовжує знижуватися, і хвороба все більше впливає на рух і фізичні можливості.

На цій стадії люди зазвичай зустрічаються із наступним:

- **Втрата здатності зв'язано спілкуватися.** Людина не може розмовляти або говорити способами, які мають сенс, хоча може іноді вимовляти слова або фрази.
- **Потрібно щоденна допомога з особистим доглядом.** Це включає в себе повну допомогу з їжею, одяганням, використанням ванної і всіма іншими щоденними завданнями по догляду за собою.
- **Значне зниження фізичних здібностей.** Людина може перестати ходити без сторонньої допомоги, не може сидіти або тримати голову без підтримки. М'язи можуть стати жорсткими, а рефлексии ненормальними. Зрештою, людина втрачає здатність ковтати і контролювати функції сечового міхура і кишечника.

1.3 Основні методи діагностики хвороби Альцгеймера

Як було сказано в розділі 1.1, діагностика хвороби Альцгеймера складається із двох рівнів [15]:

- 1) Визначення в пацієнта наявності симптомів деменції.
- 2) Знаходження специфічних для хвороби Альцгеймера артефактів або виділення в пацієнта з деменцією таких артефактів, які явно вказували б на відсутність її.

Діагностичне дослідження першого рівня:

- інтерв'ю лікаря з пацієнтом та його знайомими;
- письмове інтерв'ю пацієнта з використанням різних опитувальників:
 - 1) the Alzheimer Disease Assessment Scale–Cognitive (ADAS–Cog);
 - 2) the Mini-Mental State Examination (MMSE);
 - 3) the Clinician's Interview-Based Impression of Change (CIBIC);

4) the Neuropsychiatric Inventory (NPI).

- Фізичний та неврологічний іспит може оцінювати:
 - 1) рефлекси;
 - 2) м'язовий тонус і сила;
 - 3) можливість встати зі стільця і пройти по кімнаті;
 - 4) зір;
 - 5) слух;
 - 6) координація;
 - 7) баланс.

Діагностичне дослідження другого рівня:

- **Лабораторні тести.**

Аналізи крові можуть допомогти лікареві виключити інші можливі причини втрати пам'яті і плутанини, такі як розлад щитовидної залози або дефіцит вітамінів.

- **Візуалізація головного мозку.**

Зображення головного мозку в даний час використовуються головним чином для виявлення видимих відхилень, пов'язаних із станами, відмінними від хвороби Альцгеймера. Нові ж методи, поки що в клінічних дослідженнях, дозволяють діагностувати саме хворобу Альцгеймера.

- 1) **Магнітно-резонансна томографія (МРТ).** МРТ технологія використовує явище магнітного резонансу. Для проведення дослідження пацієнт поміщається в сильне магнітне поле. МРТ сканування абсолютно безпечне, при умові відсутності різних магнітних (металевих) імплантів.
- 2) **Комп'ютерна томографія (КТ).** КТ використовує рентген технологію, роблячи багато 2D рентген знімків об'єкту з різних ракурсів. З яких за допомогою комп'ютерних програм створюють 3D зображення. Метод не є абсолютно безпечним, так як рентген випромінювання має іонізаційні властивості, тому кількість КТ процедур має бути обмежена

і контролюватися лікарем. В даний час метод використовується головним чином для виключення пухлин, інсультів і травм голови.

3) Позитрон-емісійна томографія (ПЕТ). Під час ПЕТ сканування в кров вводиться радіоактивний індикатор низького рівня, щоб виявити особливу функцію в мозку. Рівень безпеки методу залежить від вибраного індикатору. ПЕТ візуалізація може включати в себе наступні види:

- а. ПЕТ з використанням фтородезоксіглюкози (ФДГ).** При ПЕТ із використанням ФДГ показані ділянки мозку в яких поживні речовини погано металізуються. Визначення локацій областей з поганим метаболізмом може допомогти відрізнити хворобу Альцгеймера від інших типів деменції.
- б. Сканування амілоїдного ПЕТ.** Може виміряти міру амілоїдних відкладень в мозку. Ця візуалізація в основному використовується в дослідженнях.
- с. Візуалізація Tau Pet.** Вимірює міру нейрофібрилярного ураження головного мозку, використовується тільки в дослідженнях.

1.4 Висновок до першого розділу

Зараз діагностика хвороби Альцгеймера має доволі багато проблем. Хоча більшість з них обумовлена одним – відсутністю інформації про першопричину її виникнення. Це призводить до появи ряду методів, які по непрямым ознакам мають встановити або спростувати її наявність. Деколи ці методи є достатньо прості, особливо враховуючи сучасний стан технічних методів в медицині. Хоча простота методу не може бути його недоліком, все ж, перевірка зі вставанням зі стільця, яка

була наведена в розділі 1.3, змушує задуматися над його інформативністю, особливо на ранніх стадіях хвороби. Наявні більш технологічні методи, такі як МРТ все ще не розкрили в повному обсязі свій потенціал, оскільки занадто дорогі для великої групи пацієнтів, потребують високої кваліфікації лікаря в інтерпретації результатів. Хоча останнє напевно і можна вирішити за допомогою автоматизованої системи прийняття рішення (класифікації), зважаючи на останні результати в цій сфері [16]. Тож очевидно, що розробка автоматизованої системи діагностики дозволить вивести діагностику хвороби Альцгеймера на новий рівень: зменшить імовірність людської помилки, розвантажить лікарів.

2 МАШИННЕ НАВЧАННЯ

2.1 Історичні передумови виникнення систем машинного навчання

Історично, перші образи механістичних людей і штучних істот з'являтися ще в грецьких міфах, таких як золоті роботи Гефеста і Галатея Пігмаліона. У середні віки з'явилися чутки про таємні містичні або алхімічні засоби, які здатні до розміщення розуму в матерію, такі як «Таквін», «Гомункула» і «Голема». До 19 століття ідеї про штучних людей і мислячі машини розвивалися в художній літературі: «Франкенштейн» Мері Шеллі, «Россумські універсальні роботи» Карел Чапек, і постановках, наприклад, «Дарвін серед машин» Семюеля Батлера. Штучний інтелект (ШІ) продовжує залишатися важливим елементом наукової фантастики і в наш час.

Говорячи про філософію штучного інтелекту, необхідно відмітити, що сама його концепція бере свій початок із припущення, що процес мислення людини може бути механізований. Вивчення механічного або формального мислення має давню історію. Китайські, індійські і грецькі філософи розробили структуровані методи формальної дедукції в першому тисячолітті до нашої ери. Їх ідеї розвивалися протягом століть багатьма філософами: Арістотель (дав формальний аналіз силогізму), Евклід (його «Елементи» стали зразком формального мислення), аль-Хварізм (розробив алгебру і дав своє ім'я «алгоритму»). І європейськими філософами схоластиками, такими як Вільям Оккамський, Дунс Скот та багатьма іншими.

Іспанський філософ Рамон Раймунд розробив кілька логічних машин, присвячених виробництву знань логічними засобами. Раймунд описав свої машини, як механічні об'єкти, які могли об'єднувати основні і незаперечні істини за допомогою простих логічних операцій, вироблених машиною за допомогою механічних значень, таким чином, щоб виробляти всі можливі знання.

У 17 столітті Лейбніц, Томас Гоббс і Рене Декарт досліджували можливість того, що вся раціональна думка може бути настільки ж систематична, як алгебра

або геометрія. Лейбніц пропонував універсальну мову міркувань, яка б зводила аргументацію до розрахунків. Ці філософи почали формулювати гіпотезу про систему фізичних символів, яка стане керівної вірою дослідження штучного інтелекту.

Але незважаючи на великі успіхи філософів і вчених, які займалися дослідженням штучного інтелекту, не вдавалося дати однозначного визначення для нього. Проривом в цьому стала відома стаття Алана Тьюринга в 1950 році, в якій він зазначив, що мисленню важко дати формальне визначення, а визначати його наявність в машини можна через експеримент, який потім став відомий, як «тест Тьюринга» (рис. 2.1). Суть цього тесту полягає в наступному, якщо машина могла б вести розмову через телепринтер, який не відрізнявся б від розмови з людиною, тоді було б розумним сказати, що машина думає. Ця спрощена версія проблеми дозволила Алану Тьюрингу переконливо довести, що мисляча машина була щонайменше правдоподібною. Тест Тьюринга став фундаментальним припущенням в штучному інтелекті.

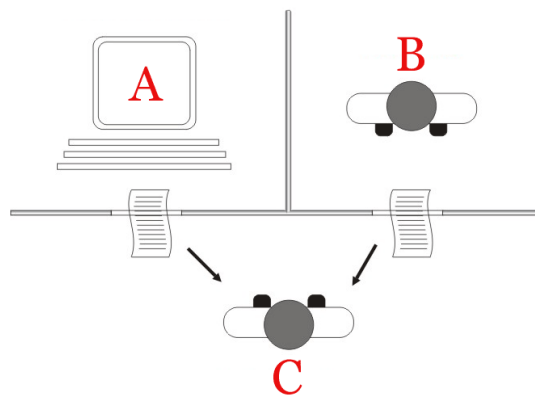


Рисунок 2.1 - Тест Тьюринга

Сама ж мисляча машина повинна була вирішувати задачі притаманні для людини:

- класифікація;
- кластеризація;
- прогнозування.

Швидко стало зрозуміло, що класичні (детерміновані) алгоритми не справляються з поставленими завданнями. Ці невдачі все сильніше підштовхували вчених до обговорення принципово нових алгоритмів, які могли б навчатися (тренуватися) в процесі роботи з інформацією. У 1959 році американський вчений Артур Семюел ввів назву «машинне навчання» для подібного класу алгоритмів.

2.2 Постановка задачі машинного навчання

Загальне формулювання задачі для алгоритмів машинного навчання виглядає наступним чином.

Є безліч об'єктів і безліч можливих результатів. Існує деяка залежність між результатами і об'єктами, але вона невідома. Відома тільки кінцева сукупність прецедентів - пар «об'єкт, результат», так звана навчальна вибірка. На основі цих даних потрібно відновити неявну залежність, тобто побудувати алгоритм, здатний для будь-якого можливого вхідного об'єкта видати досить точну відповідь. Важливою особливістю при цьому є здатність навченою системи до узагальнення, тобто до адекватного відгуку на дані, що виходять за межі навчальної вибірки. Для вимірювання точності відповідей вводиться оціночний функціонал якості, який в машинному навчанні називається – критерій якості.

Дане формулювання є узагальненням класичних задач апроксимації функцій, у яких об'єктами є вектори (рис. 2.2).

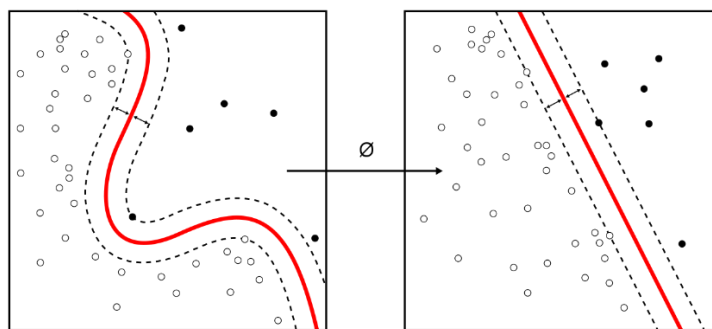


Рисунок 2.2 – Апроксимація множини на площині

У реальних задачах вхідні дані про об'єкти можуть бути неповними, неточними, нечисловими. Ці особливості призводять до великої різноманітності стратегій та алгоритмів машинного навчання.

2.3 Методи оптимізації

Існує багато методів оптимізації цільової функції. Але найбільш поширеним, у всякому разі для навчання нейронних мереж, є градієнтний метод.

Нехай існує функція залежності похибки, системи машинного навчання, від параметрів, далі будемо називати її цільовою функцією. Тоді процесом навчання буде пошук мінімуму цієї функції (рис. 2.4). Для простоти пояснення, розглянемо приклад системи з одним параметром, θ , відповідно, цільова функція буде залежати також від одного параметру. Для знаходження її мінімуму вирахуємо значення похідної в околі випадково вибраної точки, тобто випадковим чином виберемо значення параметра системи. Якщо отримане значення похідної буде додатним – це означає, що цільова функція в даному околі зростає, і навпаки. Очевидно, що для зростаючої функції - мінімум знаходиться зліва від вибраної початкової точки. Зміщуючись таким чином від попередньо вибраної точки можна знайти мінімум цільової функції. Величина цього зміщення або кроку, в загальному випадку, може варіюватися, в залежності від значення похідної. Хоча занадто велике значення кроку може приводити до того, що може бути пропущений мінімум, тобто метод не буде сходитися (рис. 2.5). Якщо буде вибраний занадто малий крок – це може призвести до великої кількості кроків для досягнення мінімуму цільової функції, і тим самим збільшить час навчання. Тому часто при виборі кроку керуються величиною похідної. Очевидно, що якщо значення похідної мале, то і функція в околі - полого, тобто навіть великий крок не призведе до того, що метод не буде сходитися.

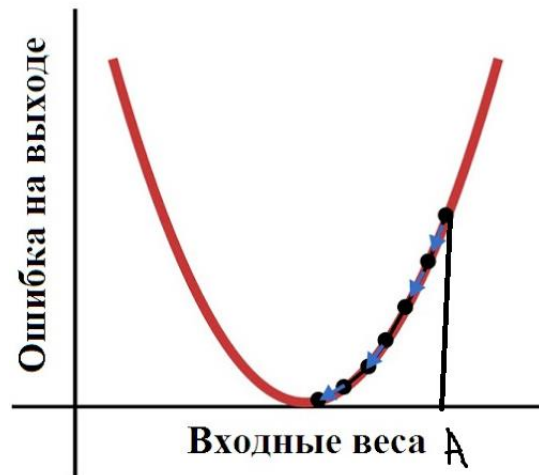


Рисунок 2.4 – Оптимізація градієнтним методом

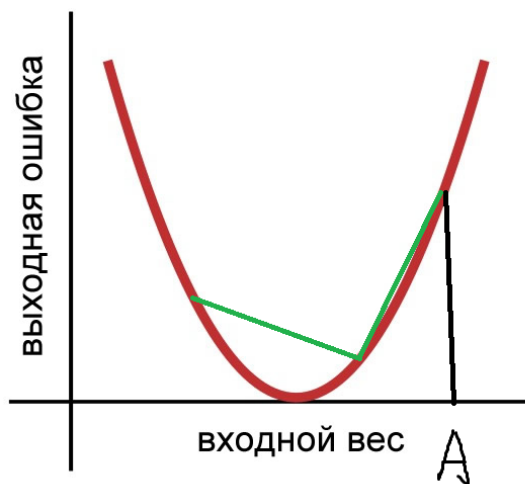


Рисунок 2.5 – Незбіжний метод

Реальні системи машинного навчання містять набагато більше, ніж один параметр. Але це не змінює методику пошуку мінімуму цільової функції, єдине що тепер потрібно враховувати - не значення похідної, а значення градієнта.

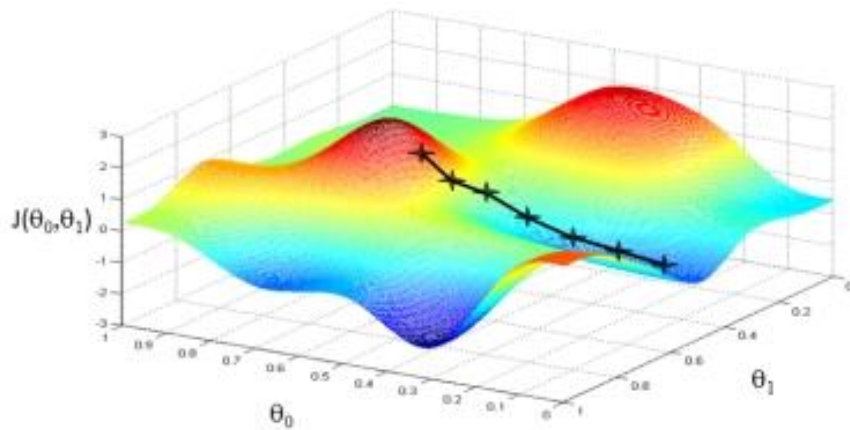


Рисунок 2.6 – Багатовимірна оптимізація

Модифікацією традиційного градієнтного метода – є метод стохастичного градієнтного спуску. Відрізняється він від класичного градієнтного методу тим, що для знаходження чергового градієнту використовується не вся навчальна вибірка, а певна її частина, сформована, в загальному, із випадково вибраних об'єктів. Це дозволяє зменшити час навчання системи, а також може зменшити імовірність перенавчання.

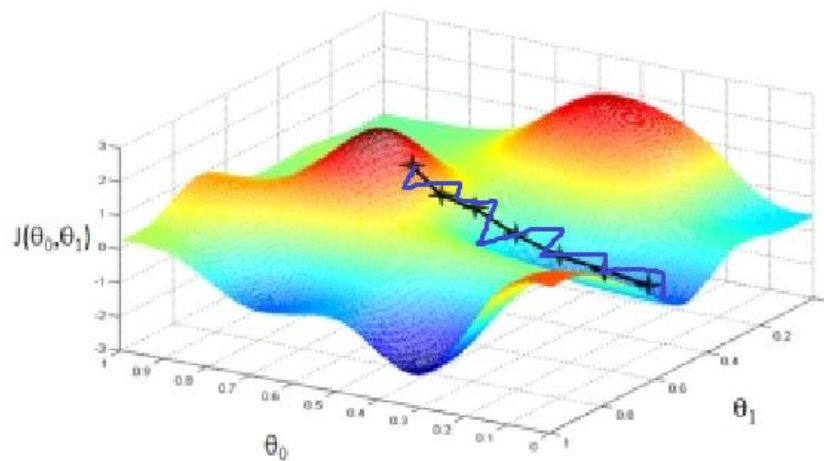


Рисунок 2.7 – Стохастичний градієнтний спуск

2.4 Проблеми машинного навчання

Неповнота об'єктів, обмежена навчальна вибірка, різномірні дані, в межах відносно невеликої вибірки, можуть призводити до виникнення ряду проблем при навчанні. Найвідомішою є проблема перенавчання (рис. 2.3). Коли апроксимуючий алгоритм, в загальному, є достатньо складним, тобто здатним абсорбувати велику кількість інформації, а навчальна вибірка досить малою або розрідженою. Тоді система може від узагальнення перейти до запам'ятовування тренувального набору. Що призводить до некоректної роботи на даних, які виходять за межі навчальної вибірки.

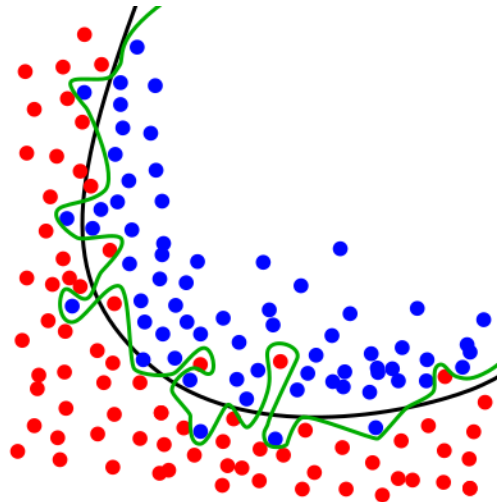


Рисунок 2.3 – Перенавчання системи машинного навчання

2.5 Алгоритми машинного навчання

Через, в тому числі, описані в розділі 2.3 проблеми, які можуть бути присутні в наборі даних, а також через специфічність в параметрах об'єктів, яка присутня в різних задачах, було розроблено ряд алгоритмів, які краще інших справляються із специфічністю даних для конкретної задачі:

- штучні нейронні мережі;
- дерева рішень;
- машини опорних векторів;
- регресійний аналіз;
- баєсовські мережі;
- генетичні алгоритми.

2.6 Стратегії машинного навчання

Завдання машинного навчання поділяють на категорії, в залежності від того в якій мірі системі, що навчається, доступний «зворотний зв'язок»:

- **Навчання з учителем.** Комп'ютеру надають приклади об'єктів і їх бажаних виходів, задані «учителем», і метою є отримання загального правила, яке буде відображає входи на виходи.
- **Напівавтоматичне навчання.** Комп'ютеру дають лише неповний тренувальний сигнал: тренувальний набір, в якому відсутні деякі (часто обчисленні) цільові виходи.
- **Активне навчання.** Комп'ютер може отримувати тренувальні мітки тільки для обмеженого набору примірників, в подальшому система може формувати запит на надання їй додаткових результатів.
- **Навчання з підкріпленням.** Тренувальні дані (у вигляді винагород і покарань) надаються тільки як зворотний зв'язок на дії програми в динамічному середовищі, як керування транспортним засобом або грі в гру з опонентом.
- **Навчання без вчителя.** Алгоритму навчання взагалі не надається міток.

Наведені стратегії навчання можуть комбінуватися між собою. А вибір оптимальної визначається класом задачі та доступність розмічених даних.

2.7 Висновок до другого розділу

Наразі методи машинного навчання є чи не єдиним кандидатом для революції в сфері діагностики хвороби Альцгеймера. Оскільки, побудова на їх основі діагностуючої системи не потребує знання причини виникнення хвороби Альцгеймера або артефактів її присутності на знімках МРТ. Що особливо важливо, враховуючи відсутність в людства знання останнього.

3 ШТУЧНІ НЕЙРОННІ МЕРЕЖІ

3.1 Сучасний стан методів глибокого навчання

Нейронні мережі на сьогоднішній день хоч і є тільки одним з багатьох методів машинного навчання, все ж, в уявленні багатьох, цілком ототожнюються з ними. І причин цьому може бути багато. По-перше, така популярність пояснюється абсолютно прекрасними результатами в задачах, які до недавнього часу вважалися підвладними тільки людині. Так в задачі класифікації зображень в 2012 році, на основі відкритої бази даних ImageNet [17], була отримана похибка 16%. Це стало ключовою подією, яка спровокувала хвилю інтересу до машинного навчання (рис. 3.1). В наступні роки похибка впала до декількох відсотків, і вже в 2014 році нейронна мережа змогла перемогти людину в точності класифікації зображень (рис. 3.2). Нейронна мережа та тодішній аспірант Стенфордського університету Андрій Карпатий змагалися в розпізнаванні більш ніж 200 порід собак.

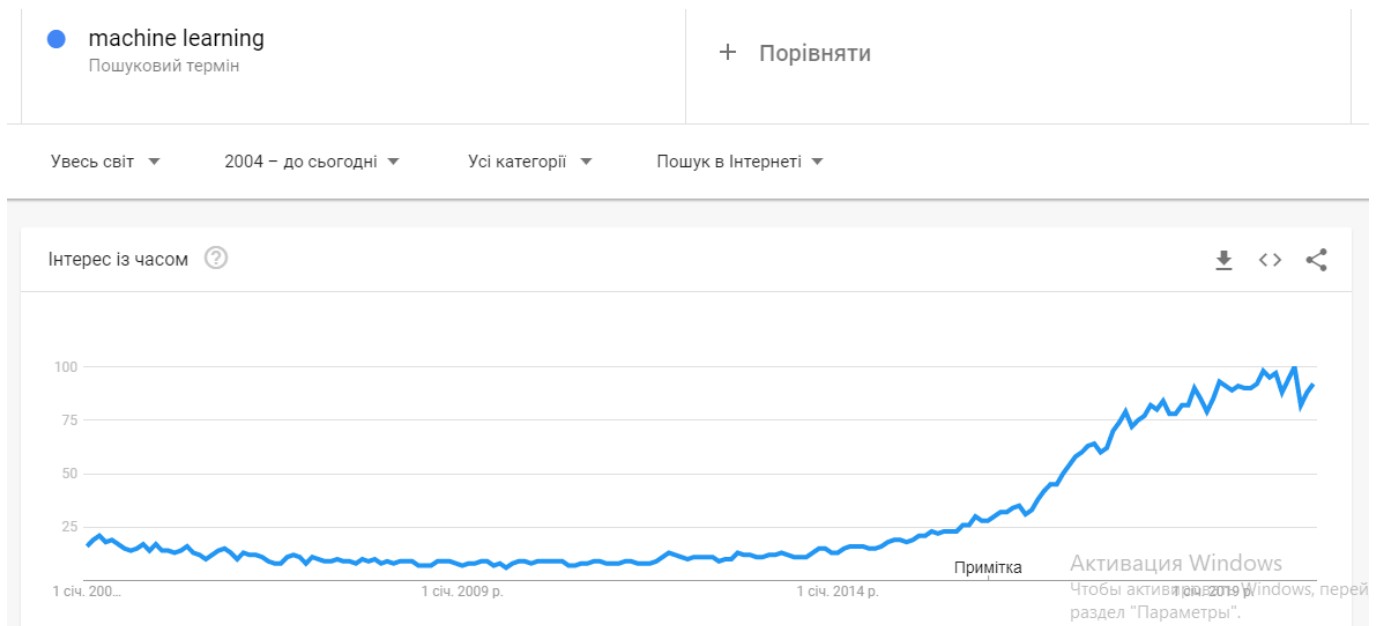


Рисунок 3.1 – Динаміка збільшення пошукових запитів за терміном “machine learning”

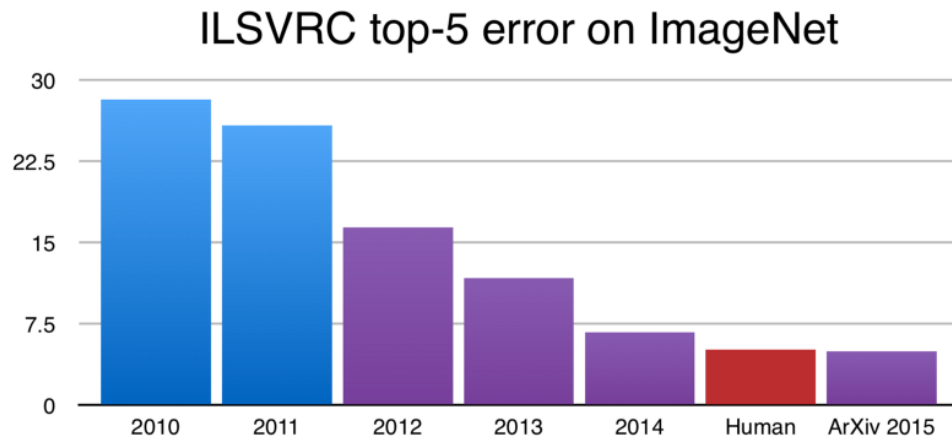


Рисунок 3.2 – Хронологія зростання точності класифікації нейронними мережами

Але класифікація зображень не є єдиною сферою де нейронні мережі отримали верх над людиною. В 2017 році програма AlphaGo, розроблена компанією Google DeepMind, перемогла в двох іграх, із запланованих трьох, найсильнішого на той час гравця в Го - Ке Цзе. Також ця програма в серії шахматних партій перемогла всі аналогічні програми, навіть ті, які отримували перемогу над чемпіонами світу.

Сфера застосування штучних нейронних мереж на сьогоднішній день надзвичайно широка:

- класифікація зображень;
- генерація тексту;
- створення музики;
- збільшення деталізації старих зображень;
- збільшення частоти дискретизації старих відеоплівків;
- сегментація зображень;
- керування автомобілем;
- перенесення стилю (рис. 3.3) та багато іншого.



Рисунок. 3.3 – Перенесення стилю

Настільки широка сфера застосування, зумовила виникнення великої кількості вакансій в ній. Так станом на травень 2020 року в США на сайті LinkedIn відкрито 34 000 вакансій в сфері Machine learning [18]. Прогнозується, що світовий ринок машинного навчання виросте із 7.3 млрд доларів в 2020 році до 30.6 млрд доларів в 2024 році [19].

3.2 Основні поняття нейронних мереж та історія їх виникнення

Наука про штучні нейронні мережі бере свій початок в 1943 році зі статті Уоррена мак-Каллока і Уолтера Піттса «Логічне числення ідей, що відносяться до нервової активності» [20, 21], в якій була запропонована модель штучного нейрона (рис. 3.4), а також описані принципи створення штучних нейронних мереж.

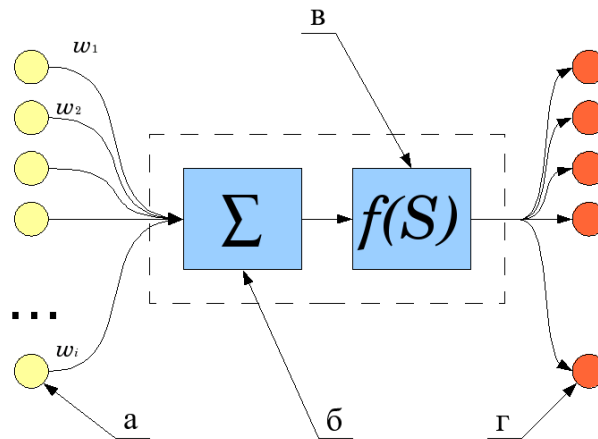


Рисунок 3.4 – Модель штучного нейрона:

а – синапси; б – суматор; в – функція активації; г – виходи; w_i – вагові коефіцієнти

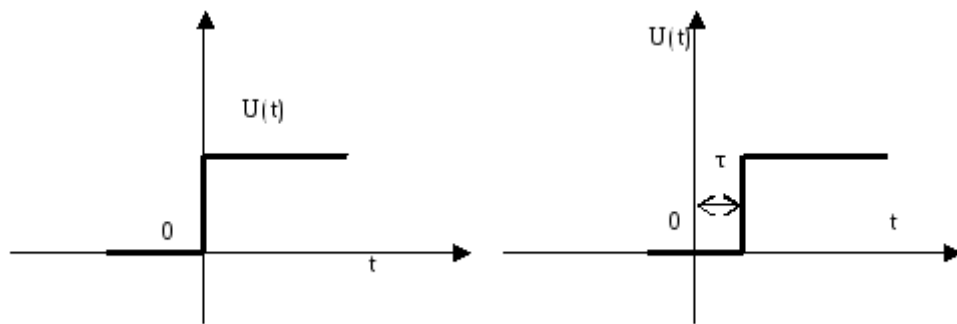


Рисунок 3.5 – Порогова функція активації:

а – не зміщена; б – зміщена на τ

Запропонована модель складалася з входів – синапсів (рис. 3.5.а), через які повинні були надходити вхідні сигнали – вектори на суматор (рис. 3.5.б). Попередньо, кожен елемент вхідного вектору повинен був перемножуватися на ваговий коефіцієнт w_i відповідного синапсу, через який цей елемент вектору поступав на суматор. Після чого отриманий результат подається на блок активації (рис. 3.4.в), після якого сигнал потрапляє на вихід (рис. 3.4.г).

Запропонована модель нейрона спочатку передбачала використання тільки порогової функції активації (рис. 3.5), швидше за все тому, що саме ця функція реалізовувала принцип «все або нічого» [22], який існує в біологічному нейроні,

але потрібно відзначити, що станом на зараз зрозуміло, що це не найкраща модель активації, і тим більше не єдино можлива. У тій же статті також були викладені теоретичні аспекти створення складних мереж на основі штучних нейронів.

Математично штучний нейрон описувався формулою:

$$y = f(u), \quad (3.1)$$

де u – індуковане локальне поле;

f – функція активації.

$$u = \sum_{i=1}^n w_i x_i + w_0 x_0, \quad (3.2)$$

де x_i – вхідні вектори;

w_i – вагові коефіцієнти;

x_i – вектор ініціалізації нейрона або вектор зміщення функції активації;

w_i – коефіцієнт ініціалізації нейрона.

Вхід ініціалізації нейрона використовується для зміщення порогової функції активації (рис. 3.1.б). Крім того, іноді до виходу нейрона спеціально додають якусь випадкову величину, яка називається зрушенням. Зрушення можна розглядати як сигнал на додатковому, завжди навантаженому вході. Зрушення використовується для зміщення порога активації. Можливі значення сигналів на входах нейрона вважають заданими в інтервалі від нуля до одиниці. Вони можуть бути дискретними (0 або 1) або аналоговими.

3.3 Порівняння біологічного і штучного нейрона

Біологічний нейрон має ряд відмінностей порівняно зі штучним:

- в штучному нейроні, на відміну від біологічного, відсутній процес розповсюдження потенціалу дії [23];
- в біологічному нейроні існує лише порогова функція активації, тоді як в штучному можуть існувати різні варіанти;
- значення на виходах біологічного нейрону можуть бути різними, тоді як значення на виходах штучного нейрона однакові.

Деякі конструктивні елементи як в біологічному нейроні, так і штучному схожі:

- наявність аксонів та дендритів в обох видах;
- наявність функції активації в обох видах;
- наявність вагових коефіцієнтів в біологічних нейронах [24].

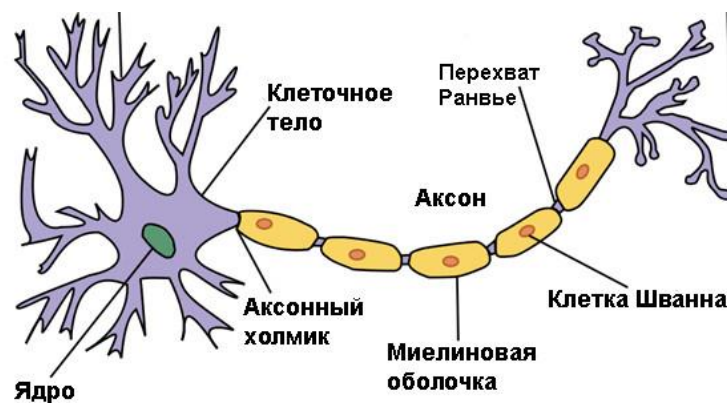


Рисунок 3.6 – Модель біологічного нейрона

3.4 Перші нейронні алгоритми (перцептрон Розенблатта)

В кінці 1940 років Дональд О. Хебб створив гіпотезу навчання, засновану на механізмі нейронної пластичності, який став відомий як Хеббове навчання. У 1954 році Фарлі і Кларк вперше використали обчислювальні машини для відтворення хеббової мережі [25].

Наступним кроком в стало створення в 1958 році американським вченим Френком Розенблатом (рис. 3.7) алгоритму, на основі штучних нейронних мереж, для класифікації зображень, який був реалізований на першому в світі нейрокомп'ютері «Марк-1» [26]. Сам алгоритм отримав назву «Перцептрон Розенблатта».



Рисунок 3.7 – Френк Розенблат з нейрокомп'ютером «Марк-1»

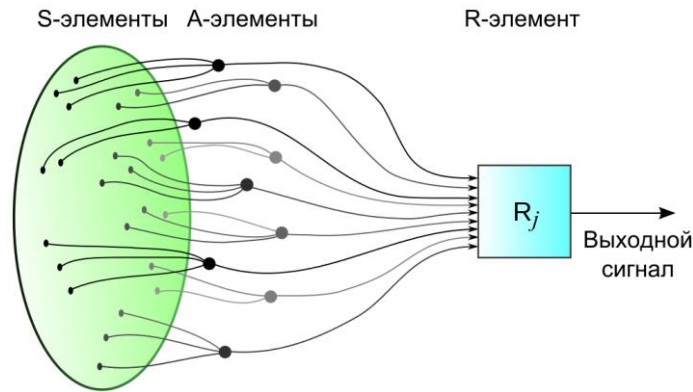


Рисунок 3.8 – Фізично еквівалентна модель перцептрона Розенблатта

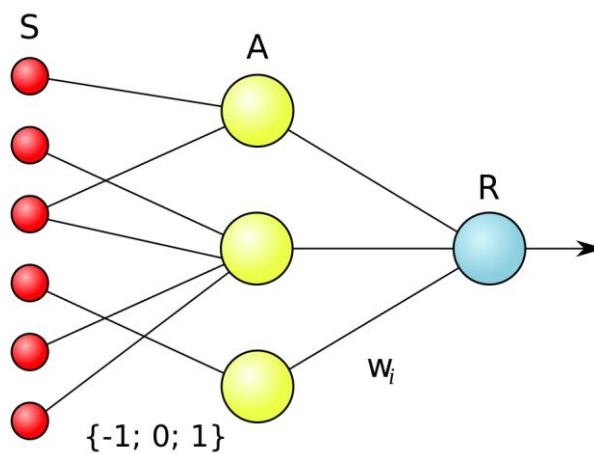


Рисунок 3.9 – Модель перцептрона Розенблатта

Елементарний перцептрон (рис. 3.8, рис.3.9) складається із трьох типів елементів: S -елементів, A -елементів та одного R -елемента. S -елементи – це шар сенсорів або рецепторів, фізичним еквівалентом є фоторезистор матриці. Кожен сенсор може знаходитися тільки в одному із двох станів – спокою або збудження, і тільки в останньому випадку він передає одиничний сигнал в наступний шар, асоціативним елементам. A -елементи називаються асоціативними, тому що кожному такому елементу, як правило відповідає цілий набір S -елементів. A -елемент активується тільки тоді, коли кількість сигналів від S -елементів, на його вході, перевищило певну величину.

Сигнали від активованих A -елементів передаються на суматор R , причому сигнал i -того асоційованого елементу попередньо перемножується на відповідний ваговий коефіцієнт w_i . Цей коефіцієнт називається вагою $A - R$ зв'язку.

Так само як і A -елементи, R -елементи підраховують суму значень вхідних сигналів. R -елемент, а разом з ним і елементарний перцептрон, видає «1», якщо сума перевищує поріг θ , інакше «-1».

Математично функцію, яка реалізується R -елементом можна записати так:

$$f(x) = \text{sign} \left(\sum_{i=1}^n w_i x_i - \theta \right), \quad (3.3)$$

де θ – поріг активації.

Навчання елементарного перцептрону полягає в зміні вагових коефіцієнтів w_i зав'язків $A - R$.

Після навчання перцептрон готовий працювати в режимі розпізнавання або узагальнення. В цьому режимі перцептрону надають раніше невідомі об'єкти, тобто такі на яких перцептрон не навчався, і перцептрон повинен встановити, до якого класу належать об'єкти. Робота перцептрона полягає в наступному: при пред'явленні об'єкта активні A -елементи передають сигнал R -елементу, якщо ця сума додатна, то приймається рішення, що даний об'єкт належить до першого класу, якщо від'ємна – то до другого.

В 1969 році Марвін Мінський публікує формальний доказ обмеженості перцептрону і показує, що він нездатний вирішувати деякі завдання (проблема «парності» і «один в блоці»), пов'язані з інваріантністю уявлень. Інтерес до нейромережових алгоритмів різко падає. Зниження інтересу до штучних нейронних мереж також була продиктована відсутністю ефективного алгоритму їх навчання.

В 1974 році Пол Дж. Вербос і Олександр Галушкін одночасно винаходять алгоритм зворотного поширення помилки (backpropagation) для навчання багатосарових перцептронів [27, 28]. Винахід не привернув особливої уваги.

У 1986 Румельхартом, Хінтон і Вільямсом, незалежно, і одночасно С. І. Барцев, та В. О. Охонін перевідкрили і суттєво розвинули метод зворотного поширення помилки [29, 30]. Це дало поштовх для початок нової ери інтересу до штучних нейронних мереж.

3.5 Метод зворотного поширення похибки

Зараз більше 99% завдань навчання нейронних мереж вирішується саме за допомогою зворотного поширення похибки [31].

Алгоритм зворотного поширення помилки належить до градієнтних методів, які розглядалися в підрозділі 2.6.

Розглянемо для початку штучну нейронну мережу, яка складається із двох шарів - перший вхідний, і на (рис. 3.9) не зображений.

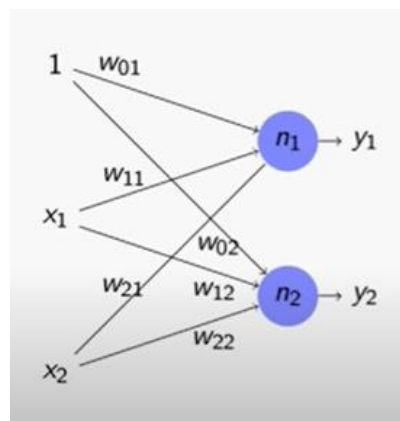


Рисунок 3.10 – Двошарова нейронна мережа

Запишемо функцію втрат штучної нейронної мережі для одного прикладу D_k , залежну від вектору правильної відповіді та вектору відповіді штучної нейронної мережі:

$$D_k(y_1, y_2, \dots, y_n) = (y_1 - a_1)^2 + (y_2 - a_2)^2 \dots + (y_n - a_n)^2, \quad (3.4)$$

де y_i – координати вектору відповіді штучної нейронної мережі;

a_i – координати вектору правильної відповіді.

Очевидно, що y_i залежить від вагових коефіцієнтів. Запишемо це в математичному вигляді так:

$$y_i = f \left(\sum_{j=0}^m x_j w_{ji} \right), \quad (3.5)$$

де w – вагові коефіцієнти;

i – номер нейрону;

j – номер синапсу i -того нейрону.

$$S_i = \sum_{j=0}^m x_j w_{ji} \quad (3.6)$$

Тоді функцію втрат штучної нейронної мережі E_k , залежну від вагових коефіцієнтів можна записати у наступному вигляді:

$$E_k(W) = D_k(y_1(w_{11}, w_{21}, \dots, w_{m1}), \dots, y_n(w_{1n}, w_{2n}, \dots, w_{mn})), \quad (3.7)$$

де W – вектор вагових коефіцієнтів;

n – номер нейрона;

m – номер синапсу n -того нейрона.

Маючи функцію втрат, залежну від вагових коефіцієнтів, можемо знайти градієнт цієї функції. Попередньо вирахувавши часткові похідні по кожному ваговому коефіцієнту. Детально цей процес описаний в розділі 2.6.

Запишемо загальну формулу для обчислення часткової похідної i -того вагового коефіцієнту використовуючи формулу складної похідної:

$$\frac{\partial E_k}{\partial w_{ji}} = \sum_{l=1}^n \frac{\partial D_k}{\partial y_l} \frac{\partial y_l}{\partial w_{ji}} \quad (3.8)$$

Знайдемо часткову похідну D_k по y_l , тоді враховуючи формулу (3.4) маємо:

$$\frac{\partial D_k}{\partial y_l} = 2(y_l - a_l), \quad (3.9)$$

Знайдемо часткову похідну y_l по w_{ji} :

$$\frac{\partial y_l}{\partial w_{ji}} = f'(S_i)x_j \quad (3.10)$$

Тоді часткова похідна E_k по w_{ji} прийме вигляд:

$$\frac{\partial E_k}{\partial w_{ji}} = 2(y_i - a_i)f'(S_i)x_j \quad (3.11)$$

Розглянемо штучну нейронну мережу, яка складається із трьох шарів – перший шар вхідний, і на (рис. 3.10) не зображений.

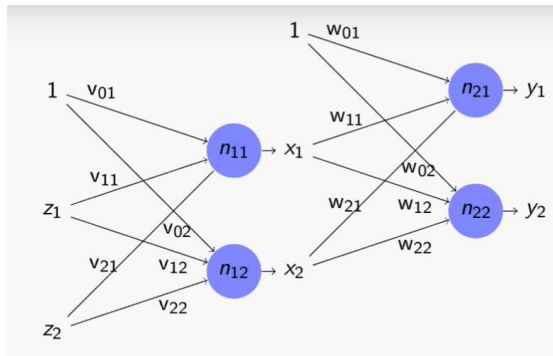


Рисунок 3.11 – Тришарова нейронна мережа

В цьому випадку E_k можна записати так само, як і для випадку двошарової штучної нейронної мережі (рис. 3.9):

$$E_k(W) = D_k(y_1, y_2, \dots, y_n), \quad (3.12)$$

Але в цьому випадку y_i залежить не від w_i , а від координат вихідного вектору першого шару, тобто x_i :

$$y_i = y_i(x_1, x_2, \dots, x_m), \quad (3.13)$$

Самі координати вихідного вектору першого шару x_j залежать від вагових коефіцієнтів v_i синапсів першого шару:

$$y_i = y_i(x_1, x_2, \dots, x_m), \quad (3.14)$$

Тоді можемо вирахувати часткову похідну y_i по x_i . Враховуючі аналогію з формулою (3.11) отримаємо:

$$\begin{aligned}\frac{\partial y_1}{\partial x_i} &= f'(S_i)w_{j1}, \\ \frac{\partial y_2}{\partial x_i} &= f'(S_i)w_{j2}, \\ &\dots\dots\dots \\ \frac{\partial y_n}{\partial x_i} &= f'(S_i)w_{jn}\end{aligned}\tag{3.15}$$

Тоді часткова похідна D_k по x_i , застосовуючи формулу похідної складної функції, буде:

$$\frac{\partial D_k}{\partial x_i} = \sum_{i=1}^n \frac{\partial D_k}{\partial y_i} \frac{\partial y_i}{\partial x_j}\tag{3.16}$$

Враховуючи попередні викладки, отримуємо:

$$\frac{\partial D_k}{\partial x_i} = 2 \sum_{i=1}^n (y_i - a_i) f'(S_i) w_{ji}\tag{3.17}$$

Таким чином отримавши це співвідношення ми можемо знайти часткову похідну функції втрат по ваговим коефіцієнтам v_i :

$$\frac{\partial E_k}{\partial v_{rs}} = \sum_{j=1}^m \frac{\partial D_k}{\partial x_j} \frac{\partial x_j}{\partial v_{rs}}\tag{3.18}$$

Так само можна продовжити і знайти часткову похідну функції втрат по входам z_l :

$$\frac{\partial D_k}{\partial z_l} = \sum_{j=1}^m \frac{\partial D_k}{\partial x_j} \frac{\partial x_j}{\partial z_l} \quad (3.19)$$

Зараз використовують дещо модифікований алгоритм зворотного розповсюдження помилки – Adam [32]. Цей алгоритм розширює backpropagation тим, що величина необхідного градієнтного кроку вибирається автоматично.

3.6 Повнозв'язані нейронні мережі

Повнозв'язана нейронна мережа (рис 3.11) – є найпростішим видом штучних нейронних мереж в якому всі нейрони наступного шару з'єднані з усіма нейронами попереднього шару. Подібний тип нейронних мереж продемонстрував хороші результати в багатьох задачах, зокрема, в задачі оцінки вартості нерухомості в Каліфорнії [33]. А також повнозв'язані шари часто використовуються, як вихідний шар згорткових нейронних мереж.

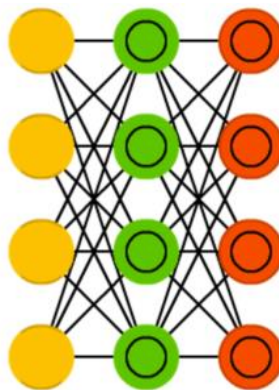


Рисунок 3.112 – Повно зв'язана нейронна мережа

Частковим випадком повно зв'язаних нейронних мереж – є мережі Гопфілда (рис. 3.12) [34]. В цій мережі виходи кожного нейрона одночасно є входами для інших нейронів.

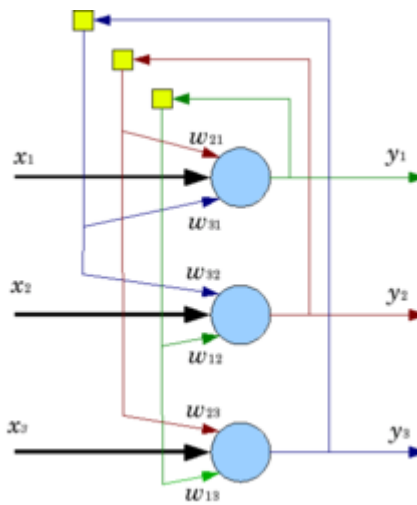


Рисунок 3.13 – Мережа Гопфілда

3.7 Згорткові нейронні мережі

Згорткова нейронна мережа або Convolutional neural network (CNN) – перші практичні результати застосування цього типу нейронних мереж з'явилися в роботі французького вченого Яна Лекуна [31].

Принцип роботи загорткової нейронної мережі полягає в тому, що згорткове ядро наче ковзає по вхідній матриці, тим самим формуючи вихід.

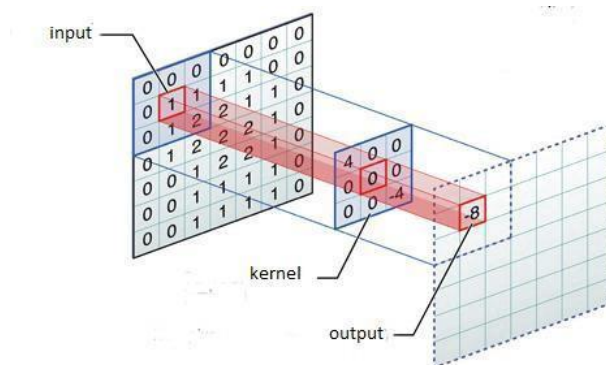


Рисунок 3.14 – Згорткова нейронна мережа

3.8 Висновок до третього розділу

Нейронні мережі з усіх алгоритмів машинного навчання показують найкращі результати в задачі класифікації зображення. Враховуючи, що система автоматичною діагностики хвороби Альцгеймера на основі знімків МРТ – це по своїй суті система класифікації зображень, то доцільно використовувати саме нейронні мережі, як основу для системи діагностики.

4 ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ В ДІАГНОСТИЦІ ХВОРОБИ АЛЬЦГЕЙМЕРА

4.1 Основні стратегії проектування систем діагностики

Базуючись на статті [35] станом на кінець лютого 2019 року виконано 71 унікальне дослідження застосування штучних нейронних мереж для діагностики хвороби Альцгеймера на основі знімків МРТ. Приблизно половина з них, а саме 32 дослідження, використовували в якості базового блока для побудови класифікатора модель CNN [31]. З огляду на те, що моделі CNN демонструють в задачах обробки зображень значно кращі результати, порівняно з повно зв'язаними моделями, доцільно розглянути саме їх.

Існуючі класифікатори на основі моделей CNN можна розділити на декілька видів:

1) **2D slice-level CNN.**

Дослідження [36-41] використовували в якості вхідних даних двовірні вектори, отримані шляхом поділу тримірного МРТ зображення по кожній осі. Перевага цього підходу полягає в тому, що існуючі CNN, які мали величезний успіх для класифікації зображень із набору ImageNet, наприклад, ResNet [42] і VGGNet [43] можуть бути легко запозичені і використані в режимі трансферного навчання. Іншою перевагою є збільшення розміру навчальної вибірки, оскільки один тримірний МРТ знімок містить десятки 2D зрізів.

2) **3D patch-level CNN.**

Цей підхід дозволяє боротися із відсутністю 3D інформації, яка виникає в методі 2D slice-level. Для нього входом слугують 3D вектори, які являють із себе частини цілого 3D МРТ зображення. Це дозволяє, як і в підході 2D slice-level, збільшити розмір вибірки, оскільки вона складатися не із 3D зображень МРТ, а із багатьох блоків МРТ зображень, самі блоки можуть перекриватися між собою. Додатковими перевагами цього методу є більш

низьке використання пам'яті, яке може бути корисне при обмежених ресурсах, і менша кількість параметрів для вивчення. Однак ця перевага присутня тільки тоді, коли використовується одна і та ж мережа для всіх блоків. Зокрема, в дослідженні [44] використовували дуже великі блоки. МРТ зображення були розділені на 27 блоків, які перекривалися. Розмір кожного складав $50 \times 41 \times 40$ пікселів, при загальному розмірі зображення $100 \times 81 \times 80$ пікселів. Було індивідуально проведене навчання для 27 CNN мереж на кожному блоці, відповідно. Самі мережі склалися із чотирьох згорткових і двох повнозв'язаних шарів. Потім був підготовлений ансамбль для прийняття рішення на основі вихідних даних 27 мереж.

Дослідження [45] використовувало менші блоки, розміром $32 \times 32 \times 32$ пікселів. При цьому була врахована можлива невідповідність між ділянками, отриманими за однаковими координатами з різних об'єктів. Щоб уникнути цієї відмінності між об'єктами без застосування методів нелінійної реєстрації, блоки збиралися за допомогою методу k -середніх. У наступних дослідженнях [46, 47], використовувалися ще більш дрібні блоки, розміром $19 \times 19 \times 19$. Підхід [47] трохи відрізняється, так як в ньому вважалося, що базові блоки системи не можуть бути навчені окремо. Тобто для кожного блоку розробляється своя система.

3) ROI-based CNN.

Методи patch-level використовують всю МРТ, розбиваючи її на блоки. Тим не менше, більшість блоків не інформативні, так як вони містять частини мозку, які не схильні до хвороби Альцгеймера. Методи, засновані на областях інтересу, долають цю проблему, зосереджуючись на регіонах, які, як відомо, є інформативними. Таким чином, складність структури мережі може бути зменшена, оскільки використовуються менш зашумлені дані. В дослідженнях [48, 49] областю інтересу був гіпокамп. Дослідження відрізняються визначенням області інтересу гіпокампу.

4) 3D subject-level CNN.

В деяких дослідженнях використовувався 3D-підхід на рівні об'єкта. При такому підході вся МРТ використовується одночасно і класифікація здійснюється на рівні об'єкта. Перевага полягає в тому, що просторова інформація повністю інтегрована. У деяких дослідженнях було адаптовано дві класичні архітектури: ResNet [42] і VGGNet [43]. Також в ряді досліджень було запропоновано ряд оригінальних архітектур [50-52]. Ці методи не демонструють стабільності. Основною перешкодою для їх успішної реалізації є недостатня кількість об'єктів, порівняно з попередніми методами. Ця різниця складає сотні, а то і тисячі разів. Недостатня кількість об'єктів є умовою перенавчання нейронної мережі.

4.2 Переваги методу 2D slice-level CNN

Найкращу точність демонструють 2D slice-level CNN методи. Це може бути пов'язано з багатьма причинами:

- **Великий розмір навчальної вибірки.** Всього з одного МРТ зображення розміром $100 \times 100 \times 100$ вдається отримати 300 об'єктів, і тим самим уникнути перенавчання.
- **Застосування трансферного навчання.** Дозволяє використовувати найкращі приклади попередньо навчених моделей, таких як ResNet [42], VGGNet [43] і так далі.
- **Близькі до ідеалу архітектури.** Вони були неодноразово протестовані в різних задачах і показали високу ефективність.
- **Легкість використання 2D slice-level CNN моделей штучних нейронних мереж.** Моделі для 2D slice-level CNN методу вибираються, як правило, із переліку готових моделей, які, в тому числі, присутні в якості готового блоку в фреймворку tensorflow. Це дозволяє

сконцентруватися саме на експерименті, перебираючи різні варіанти його проведення.

Таблиця 4.1 – Точність систем діагностики хвороби Альцгеймера на основі штучних нейронних мереж [53]

Назва нейронної мережі	Точність
DemNet	91.85%
Gupta et al. (Natural Bases)	85.00%
Gupta et al. (MRI Bases)	78.20%
Payan and Montana (2D Convolutions)	85.53%
Payan and Montana (3D Convolutions)	89.47%

4.3 Проблеми в автоматичній класифікації знімків МРТ

В глибокому навчання і в класифікації знімків МРТ, зокрема, існує проблема з відповідністю результатів, яка може зумовлюватися наступними причинами [54]:

- Відсутністю єдиної тестової бази даних.
- Відсутність загально прийнятого алгоритму тестування. Із-за використання різних метрик прямо порівнювати результати, які були отримані за допомогою 2D slice-level CNN и 3D patch-level CNN не завжди коректно.
- Різна степінь попередньої обробки даних суттєво впливає на результат. Тому залишається не зрозумілим, якщо в дослідженні була використана нова архітектура, в якій степні отриманий результат завдячує архітектурі, а в якій обробці даних.
- Відсутність прозорої процедури опису всіх етапів створення діагностуючої системи є фактором, який не дозволяє об'єктивно

порівнювати отримані результати. Це особливо критично в експерименті з використання методів глибокого навчання, коли із-за неповного опису вхідних даних часто буває неможливо його відтворити.

- Також автори часто в висновках вказують тільки найбільш загальні показники точності, залишаючи не освітленими показники стабільності системи.

4.4 Висновок до четвертого розділу

Використання нейронних мереж в діагностиці хвороби Альцгеймера має, як для сфери штучних нейронних мереж, доволі велику історію. На протязі якої було розроблено ряд моделей та методів навчання. Отримані результати демонструють осяжність перспективи побудови надійної автоматичної системи діагностики. Хоча існують проблеми, зокрема, невелика кількість даних, невідповідність результатів для порівняння алгоритмів.

5 ОПИС РОЗРОБЛЕНОЇ СИСТЕМИ

5.1 Опис архітектури

За основу для побудови системи автоматичної діагностики хвороби Альцгеймера була обрана неймережева модель ResNet50V2.

Дана модель була обрана виходячи із наступних засад:

- обмеженість обчислювальних ресурсів та жорсткі часові рамки виконання дипломної роботи не дозволяють проводити широкомасштабні експерименти по розробці принципово нового дизайну нейронної мережі;
- наявність подібних існуючих систем давала впевненість в результативності;
- наявність підготовлених коефіцієнтів, в значній мірі нівелювала проблему можливого закінчення експерименту з нульовим результатом, а також зменшувала обчислювальну складність.

Модель була пристосована для діагностики трьох класів МРТ зображень, шляхом заміни останнього її шару, на повнозв'язаний шар, який складається з трьох нейронів.

В якості оптимізатора використовувався алгоритм Adam [32] із пакету tensorflow – версія 2.2.0 [55].

Функцією втрат виступила «categorical_crossentropy»:

$$H(p, h) = - \sum_{i=0}^N p_i \log h_i, \quad (5.1)$$

де p – прогноз нейронної мережі;

h - правильна відповідь.

5.2 Опис даних та точності розробленої системи

Для навчання системи була використана база даних, сформована на основі відкритої бази даних ADNI [56]. Всього було підготовлено 673 3D МРТ зображень. Всі об'єкти із бази даних належать до трьох класів: Healthy Controls – 187 зображень, Mild Cognitive Impairment – 326 зображень, Alzheimer's – 160 зображень. Кожне зображення мало форму 89 x 89 x 89 пікселів. Подібна диспропорція між величиною класів не є критичною, так як Mild Cognitive Impairment являє найбільшу діагностичну цінність. В подальшому всі 3D зображення були на розділені на 2D зрізи по кожній осі. Наступним кроком став розподіл сформованої бази даних на тренувальну та тестову частину. Тренувальна частина склала 90%, тестова – 10%, відповідно.

При навчанні використовувалися зрізи починаючи з 30 до 65, виходячи з того, що саме вони містять найменше шуму, хоча б враховуючи той факт, що на них більша частина зайнята саме мозком, а не навколишнім середовищем.

Було отримано 70567 2D зображення: тренувальних – 63500; тестових -7067.

Всі отримані зрізи були приведені до розміру 86 x 86 x 86, шляхом видалення пікселів. Після цього масштабовані до розміру 224 x 224 за допомогою функції `resize` (бібліотека OpenCV – версія 4.2.0.32 в Python), використовуючи параметри за замовчанням [57]. Вектор правильних відповідей був представлений в форматі `one_hot_encoding`.

Для визначення функціональних можливостей розробленої системи була проведена серія експериментів.

Експеримент 1. Обчислення стандартних метрик: `accuracy`, `precision`, `recall`, `f`-міра на основі `confusion matrix`.

Золотим стандартом в оцінці якості класифікаторів, який дозволяє порівнювати алгоритми між собою, вважається матриця помилок (англ. `Confusion Matrix`), на основі якої обчислюються такі метрики як загальна точність (англ.

ассурасу), точність для окремого класу (англ. precision), повнота (англ. recall), f-міра.

Таблиця 5.1 – Матриця помилок

		Експертна оцінка	
		$y = 1$	$y = 0$
Оцінка системи	$y' = 1$	True Positive (TP)	False Positive (FP)
	$y' = 0$	False Negative (FN)	True Negative (TN)

Загальна точність є відношенням загальної кількості правильних відповідей, наданих класифікатором, до загальної кількості відповідей, і визначається по формулі:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.1)$$

Ця метрика погана, коли є задача з нерівним класами або коли велика загальна кількість класів.

Для оцінки якості роботи алгоритму на кожному з класів використовуються точність для окремого класу (англ. precision) та повнота (англ. recall).

$$precision = \frac{TP}{TP + FP} \quad (5.2)$$

$$recall = \frac{TP}{TP + FN} \quad (5.3)$$

Точність для кожного класу можна інтерпретувати як відсоток об'єктів, названих класифікатором позитивними і при цьому вони дійсно є позитивними. Повнота показує, яку частину позитивного класу із всіх об'єктів позитивного класу знайшов алгоритм.

Метрикою, яка узагальнює дві попередні, є f-міра:

$$f = \frac{precision * recall}{precision + recall} \quad (5.4)$$

В загальному матриця помилок це квадратна матриця розміром $N \times N$, де N – кількість класів. В ній загальна точність, точність для кожного класу, повнота та f-міра обчислюються по наступних формулах:

$$accuracy = \frac{\sum_{i=1}^n A_{i,i}}{\sum_{i=1}^n \sum_{j=1}^n A_{i,j}}, \quad (5.5)$$

де A – матриця помилок.

$$precision_c = \frac{A_{c,c}}{\sum_{i=1}^n A_{c,i}} \quad (5.6)$$

$$recall_c = \frac{A_{c,c}}{\sum_{i=1}^n A_{c,i}} \quad (5.7)$$

Таблиця 5.2 – Матриця помилок для тестової вибірки

		Експертна оцінка		
		Healthy Controls	MCI	Alzheimer's
Оцінка системи	Healthy Controls	1613	437	43
	MCI	332	2926	172
	Alzheimer's	19	60	1465

Таблиця 5.3 – Основні параметри розробленої системи

Ассурасу (для тренувальної вибірки)	99.95%
Ассурасу (для тестової вибірки)	84.95%
Precision	0.8493
Recall	0.8575
F-міра	0.8534
Precision (Healthy Control)	82.12%
Precision (MCI)	85.48%
Precision (Alzheimer's)	87.20%

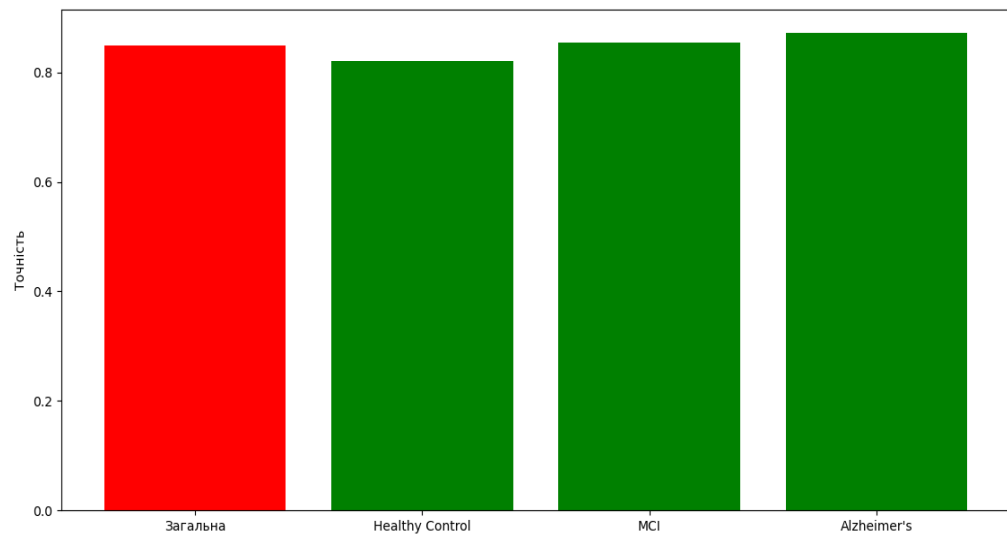


Рисунок 5.1 – Тестова точність для кожного класу

Експеримент 2. Визначення загальної точності на тестовій вибірці для кожної орієнтації зрізів.

Окремо для трьох груп об'єктів, отриманих за допомогою різних орієнтацій зрізів (vertical, horizontal, transverse) було складено три матриці помилок, відповідно. Для кожної такої групи, використовуючи відповідну їй матрицю помилок була обрахована загальна точність по формулі (5.5).

Таблиця 5.4 – Результати другого експерименту

(рис. 5.3)	(Тестова вибірка)	Точність, %
	Accuracy (загальна)	84.95
	Accuracy (vertical)	86.26
	Accuracy (horizontal)	81.41
	Accuracy (transverse)	87.16

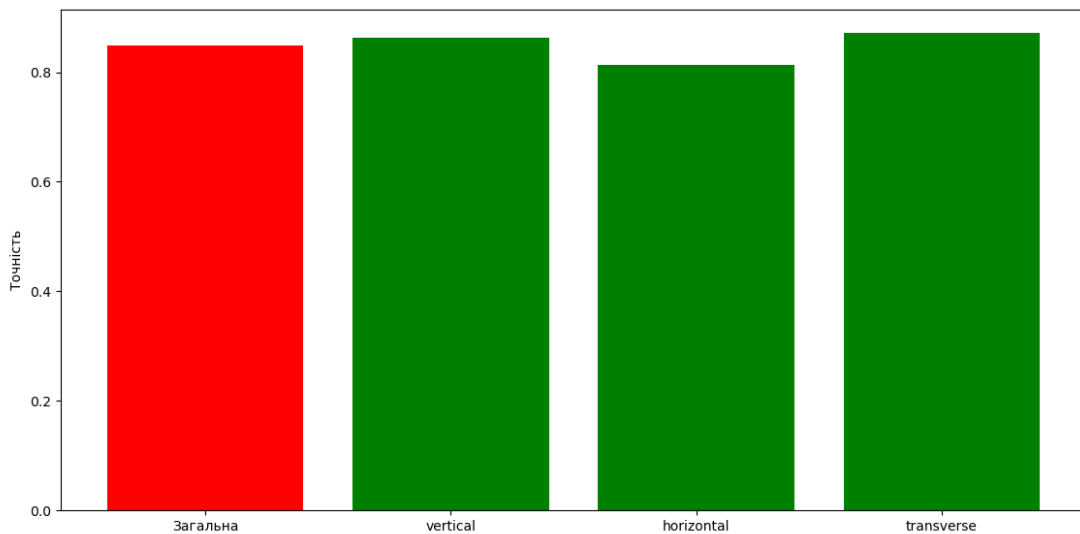


Рисунок 5.2 – Тестова точність для кожної орієнтації зрізів

Експеримент 3. Визначення загальної точності на тестовій вибірці для кожної групи об’єктів з унікальною комбінацією належності до класу та зрізу.

Експеримент 3 розширює перший та другий експеримент тим, що точність визначається не для групи об’єктів, котрі належать до одного класу чи отримані за допомогою конкретної орієнтації зрізів. А для групи об’єктів, які мають унікальну комбінацію належності до певного класу і, одночасно, були отримані за допомогою конкретної орієнтації зрізу, наприклад, група об’єктів, які належать до першого класу і були отримані за допомогою vertical зрізу чи об’єкти, які належать до третього класу і були отримані за допомогою horizontal зрізу. Тобто як і в другому експерименті було складено три матриці помилок для трьох груп об’єктів, отриманих за допомогою різних орієнтацій зрізів (vertical, horizontal, transverse), відповідно. Для кожної такої матриці обраховувалася точність для кожного класу по формулі (5.5).

Таблиця 5.5 – Результати першого експерименту

(рис. 5.4)	(Тренувальна вибірка)	Точність, %
	Загальна	99.37
	1-ший клас (vertical)	83.75
	1-ший клас (horizontal)	77.74
	1-ший клас (transverse)	84.62
	2-гий клас (vertical)	86.41
	2-гий клас (horizontal)	82.31
	2-гий клас (transverse)	87.75
	3-тій клас (vertical)	89.05
	3-тій клас (horizontal)	83.65
	3-тій клас (transverse)	88.98

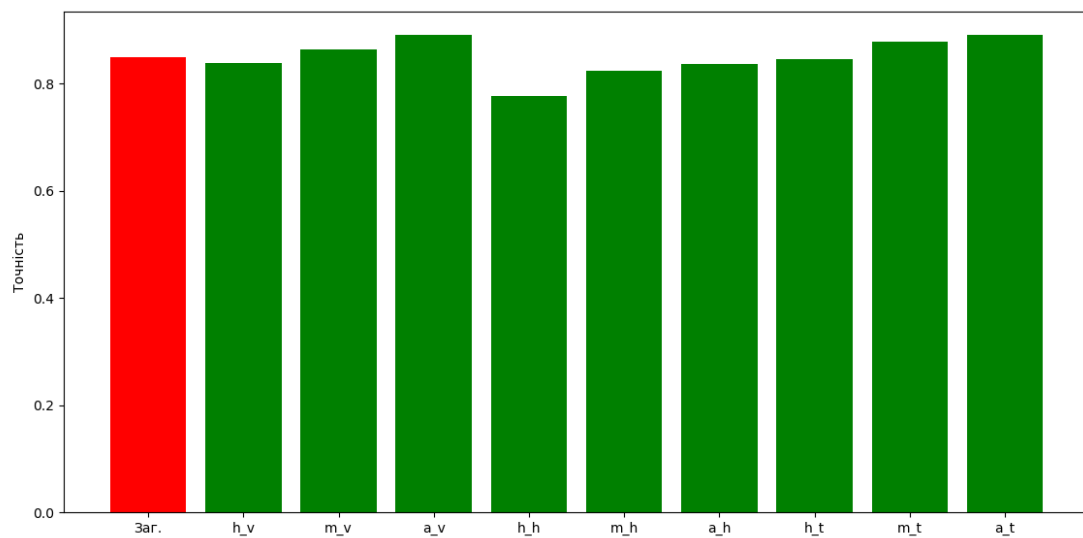


Рисунок 5.3

Експеримент 4. Визначення точності в залежності від координати МРТ зображення.

Була визначена точність для групи об'єктів (зрізів), які мають однаковий номер зрізу, не залежно від того по якій координаті із трьох (vertical, horizontal, transverse) він був отриманий.

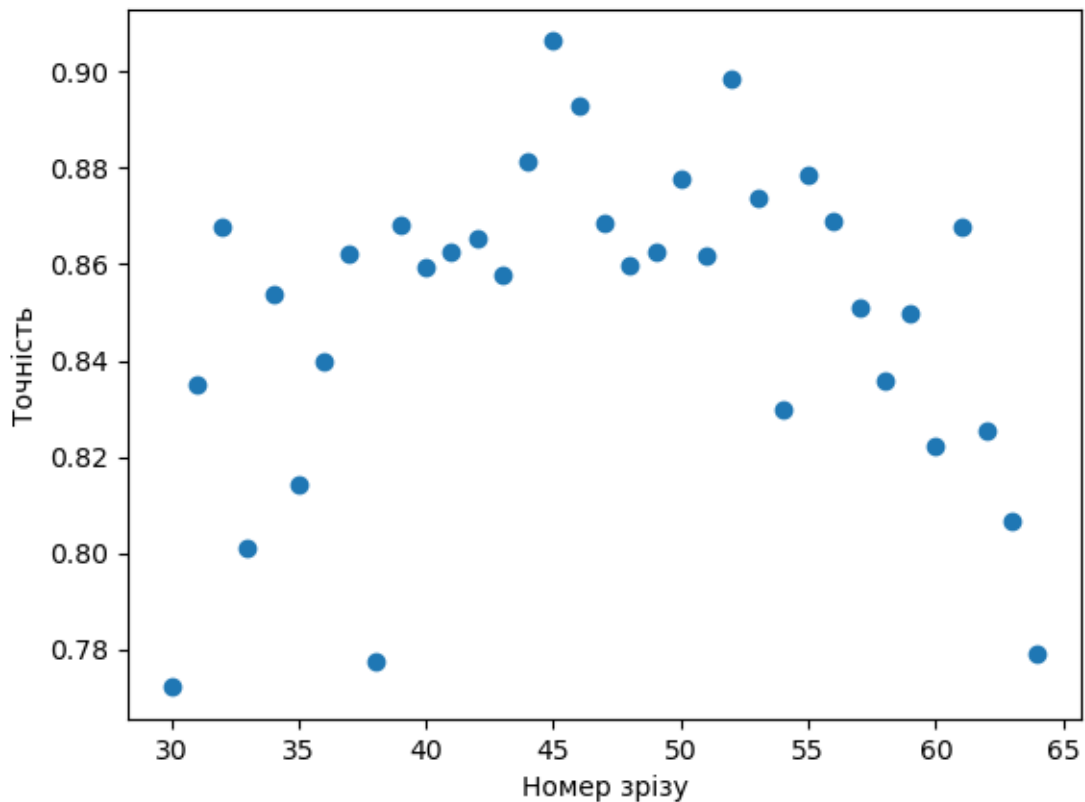


Рисунок 5.4

Експеримент 5. Визначення точності в залежності від координати зрізу, окремо, для кожної орієнтації зрізів.

Експеримент 5 розширює четвертий експеримент тим, що залежність точності від номеру зрізу була обрахована, окремо, для кожної орієнтації зрізів.

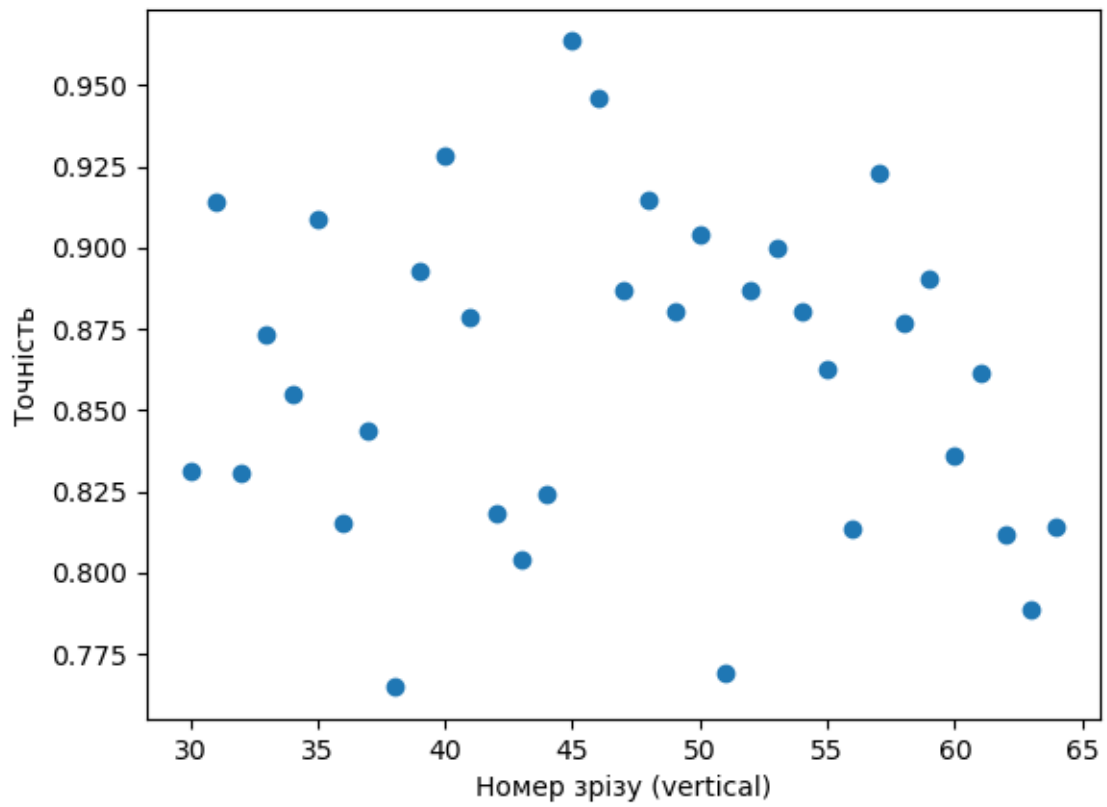


Рисунок 5.5 – Точність, в залежності від номеру зрізу (vertical)

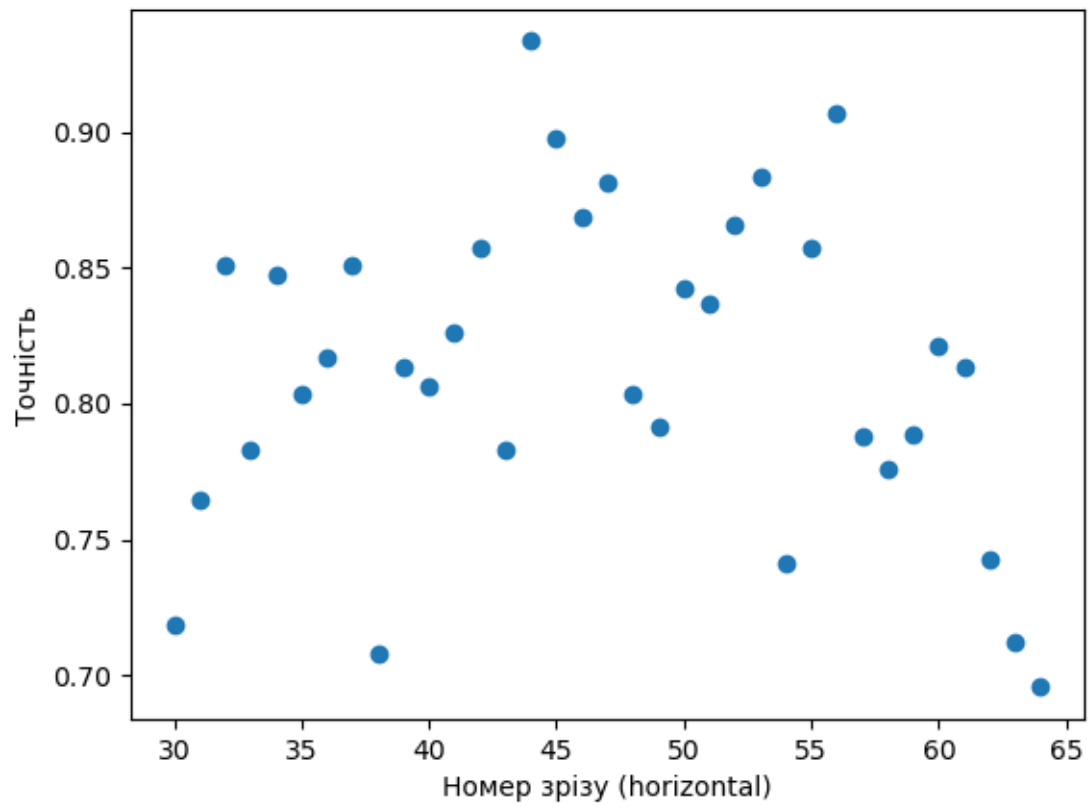


Рисунок 5.6 – Точність, в залежності від номеру зрізу (horizontal)

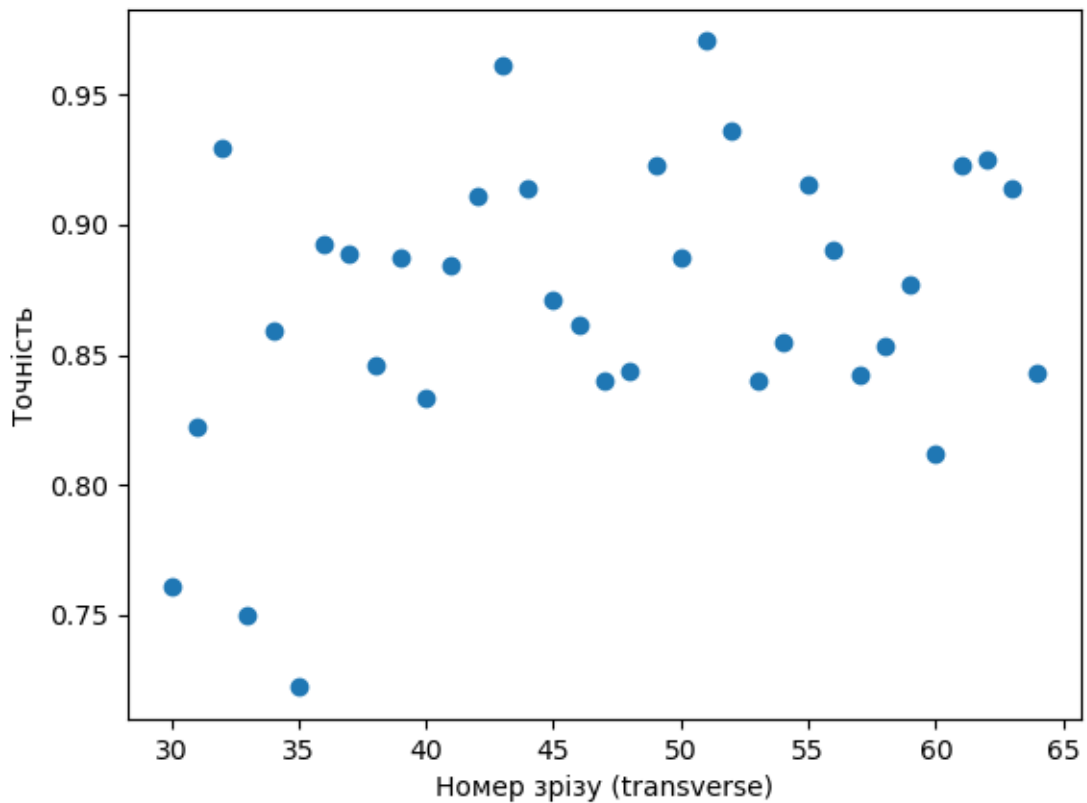


Рисунок 5.7 – Точність, в залежності від номеру зрізу (transverse)

Експеримент 6. Визначення співвідношення між координатами вихідного вектору.

Метою експерименту було отримання прихованої інформації, тобто такої, яка втрачалася в попередніх експериментах при оцінці точності, оскільки в них використовувалися логічні функції, то правильною відповіддю нейронної мережі вважалася тільки така, при якій координата найбільшого елемента вихідного вектору, співпадала з координатою елемента вектору-позначки в якому була записана одиниця. Очевидно, що для реальної медичної практики отримання неправильної відповіді від нейронної мережі може бути не критичним. Так як існує ряд ситуацій, прийняття рішення про постановку діагнозу в яких може прийматися враховуючи додаткову інформацію. Розглянемо наступну ситуацію. Існує два пацієнта, які пройшли МРТ тестування доволі давно, по причині підозри хвороби Альцгеймера, настільки давно, що хвороба Альцгеймера, якщо вона була присутня

в вигляді незначних симптомів, змогла б проявитися і бути чітко визначеною. Тобто відомо на якому знімку МРТ присутня хвороба Альцгеймера, а на якому ні. Будемо вважати, що в першого пацієнту на знімках МРТ були присутні артефакти по'язані з хворобою Альцгеймера; в другого ні. Закодуємо ці данні в вектори-позначки, причому зробимо це в форматі one_hot. Так в першого пацієнта на момент проходження дослідження існував легкий когнітивний розлад через хворобу Альцгеймера, тоді вектор позначка для нього прийме вигляд $\{0; 1; 0\}$, в іншого таких артефактів не існувало – $\{1; 0; 0\}$. Знімки обох пацієнтів були передані для аналізу автоматичній системі класифікації, і отримані наступні результати: $\{0.45; 0.4; 0.15\}$ – для першого пацієнту; $\{0.9; 0.05; 0.05\}$ – для другого пацієнта. Якщо використовувати для оцінки точності діагностуючої системи такий самий підхід, як і в попередніх експериментах, то це означало, що середня точність системи склала 50%, тобто один об'єкт правильно класифікований, інший ні. Але якщо оцінити в двох векторах співвідношення елементу з найбільшим значенням до суми двох інших, то стане видно, що воно кардинально відрізняється. Так векторі відповіді, коли система дала правильний результат, найбільший його елемент, тобто 0.9, в 9 разів більший за суму двох останніх. В випадку коли система надала не парильну відповідь найбільший елемент, тобто 0.45, в 1.2 рази менший за суму двох інших елементів. В цьому випадку можна сказати, що система сумнівається в відповіді, яку вона надала. Враховуючи вище наведене, доцільно обрахувати величину цих співвідношень.

Математично це можна записати так:

$$\begin{aligned} s_1 &= \frac{\left(\sum_{i=1}^m \frac{a_1}{a_1 + a_2 + a_3}\right)}{m}, \\ s_2 &= \frac{\left(\sum_{i=1}^m \frac{a_2}{a_1 + a_2 + a_3}\right)}{m}, \\ s_3 &= \frac{\left(\sum_{i=1}^m \frac{a_3}{a_1 + a_2 + a_3}\right)}{m} \end{aligned} \quad (5.8)$$

де s_i – середня відсоткова частина i -того елемента вектору-результату, отримана на групі об’єктів;

a_i – i -тий елемент вектору-результату.

Синім на рисунках 5.8-5.10 зображено загальне співвідношення всіх відповідей нейронної мережі. Червоним – елемент вектору відповіді, який мав вказувати на правильний результат.

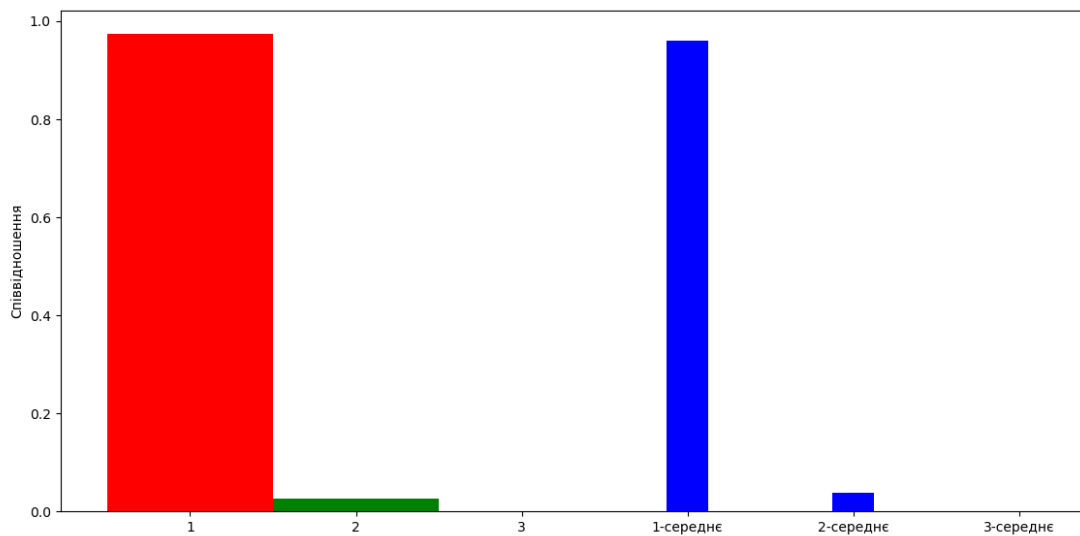


Рисунок 5.8 – Співвідношення між елементами вектору-результату, коли найбільший елемент вектору-результату (позначений червоним) вказує на правильну відповідь.

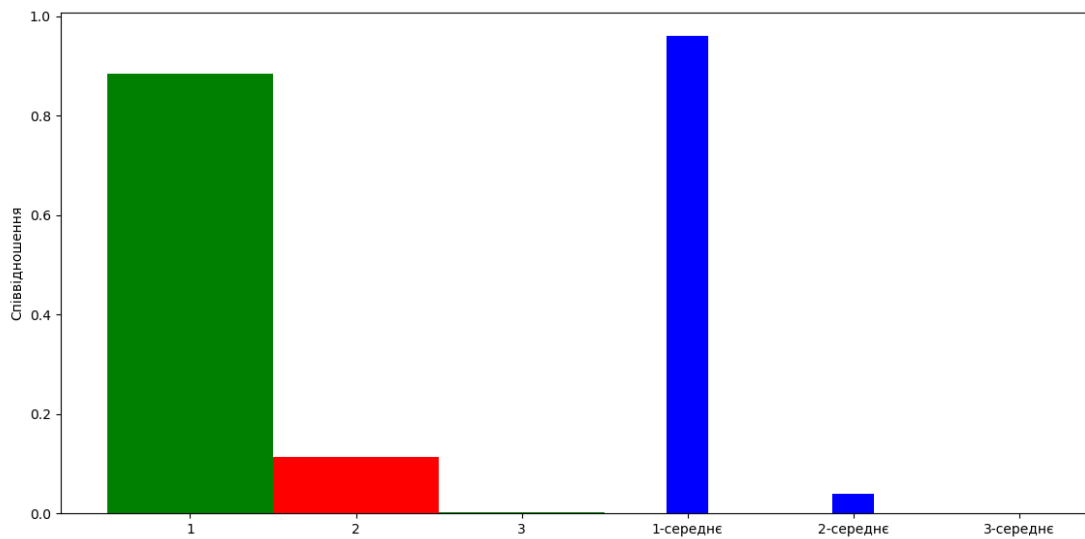


Рисунок 5.9 – Співвідношення між елементами вектору-результату, коли другий найбільший елемент вектору-результату (позначений червоним) вказує на правильну відповідь.

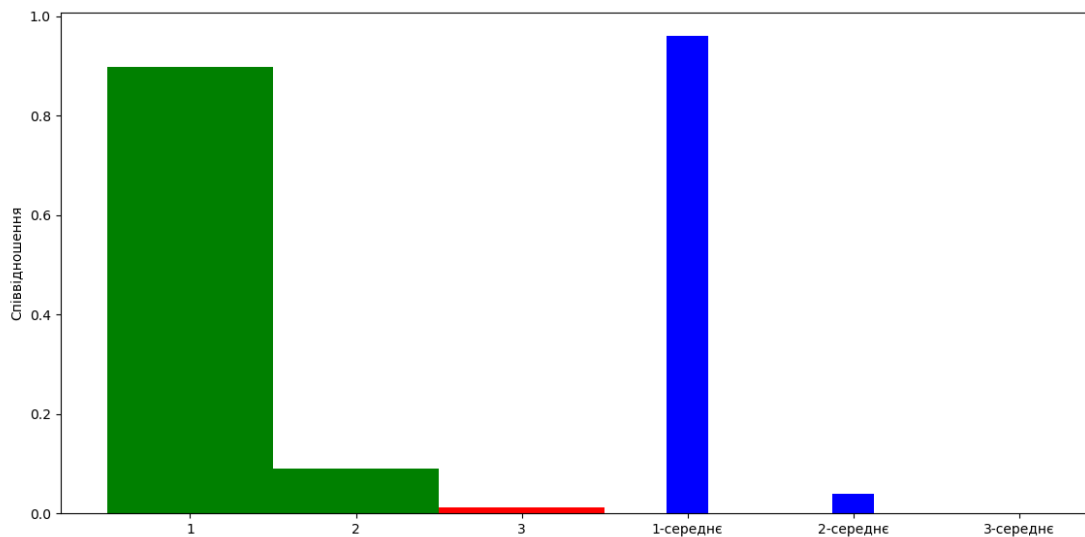


Рисунок 5.10 – Співвідношення між елементами вектору-результату, коли найменший елемент вектору-результату (позначений червоним) вказує на правильну відповідь.

Експеримент 7. Визначення точності на рівні об'єктів.

До цього точність визначалася на рівні 2D зрізів. Очевидно, що в реальній медичній практиці класифікація буде проводитися не на рівні окремих зрізів, а на рівні цілого 3D зображення, тобто одночасно всіх 2D зрізів. Оскільки в тестовій вибірці для різних об'єктів може знаходитися різна кількість зрізів, всі об'єкти були розташовані зліва на право, в порядку збільшення кількості зрізів в них. Для кожного об'єкту було обраховане відношення елементу вектору-результату, який вказує на правильну відповідь (позначено червоним) до суми інших елементів (позначено синім). Верхній графік позначає кількість зрізів в об'єкті, нижній зазначені вище співвідношення між правильною і неправильною відповіддю. Всього було сформовано три групи таких графіків, враховуючи те, що одночасно всі 673 об'єкта не поміщалися на один графік.

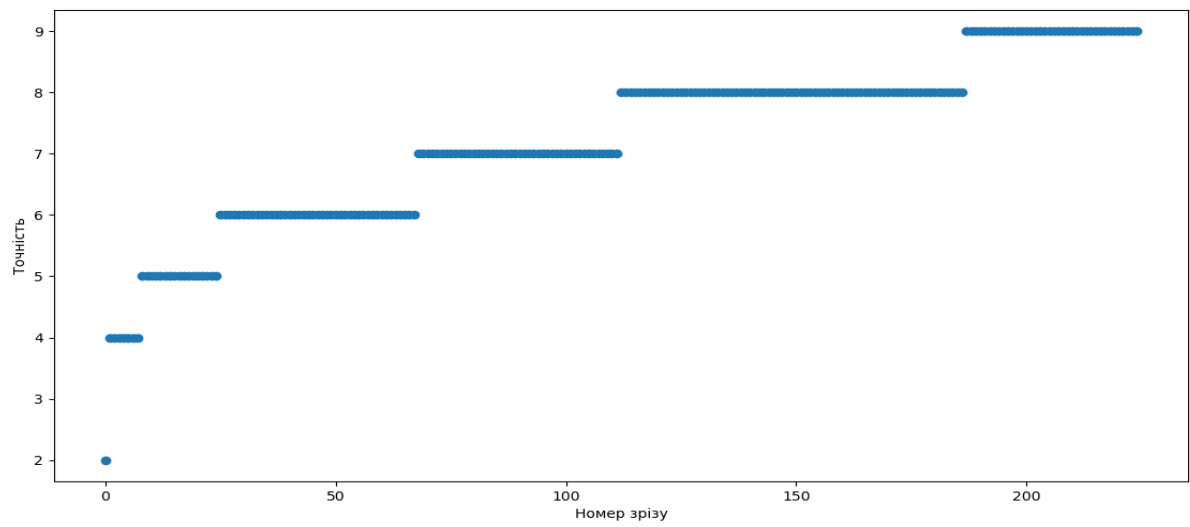


Рисунок 5.11 – Кількість зрізів в об'єктах

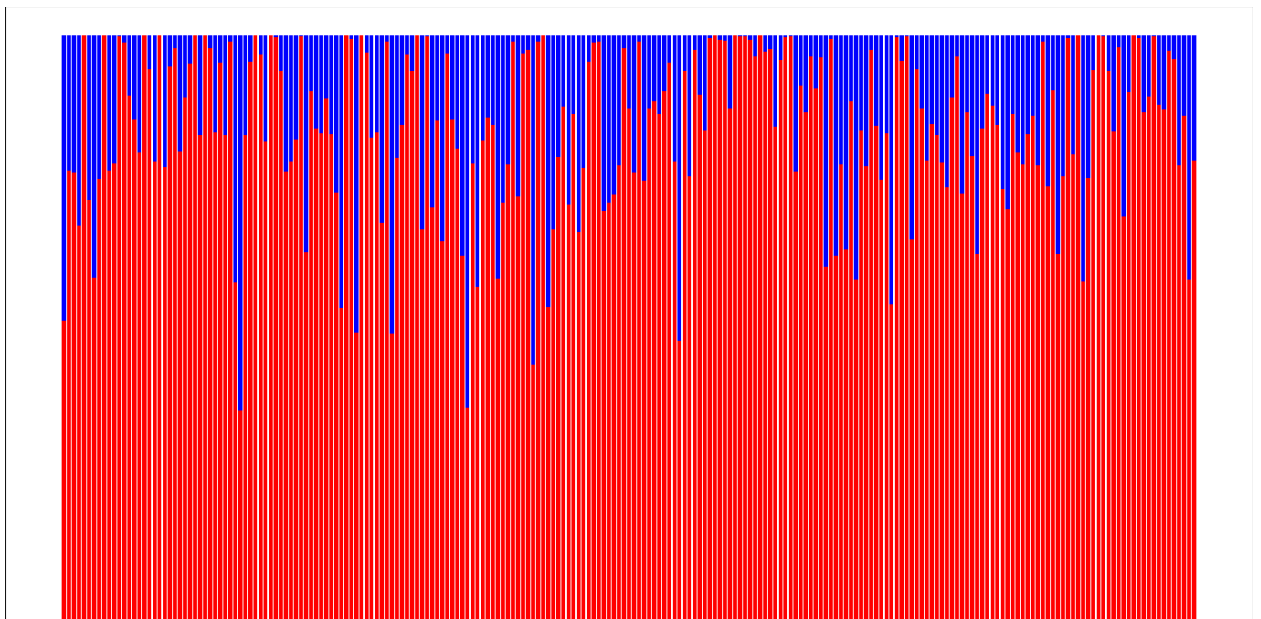


Рисунок 5.12 – Точність отримана на рівні об'єктів

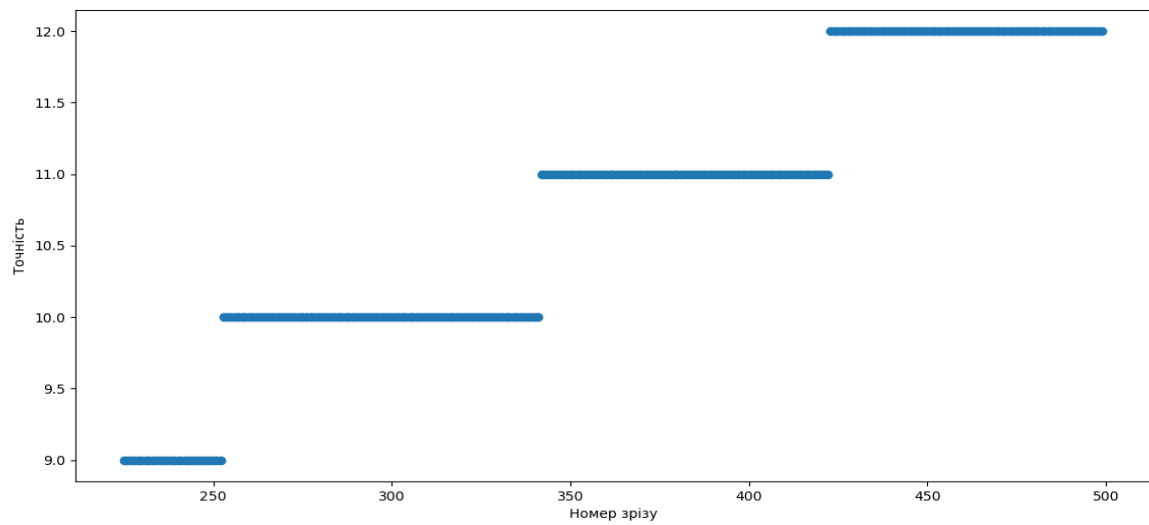


Рисунок 5.13 – Кількість зрізів в об'єктах

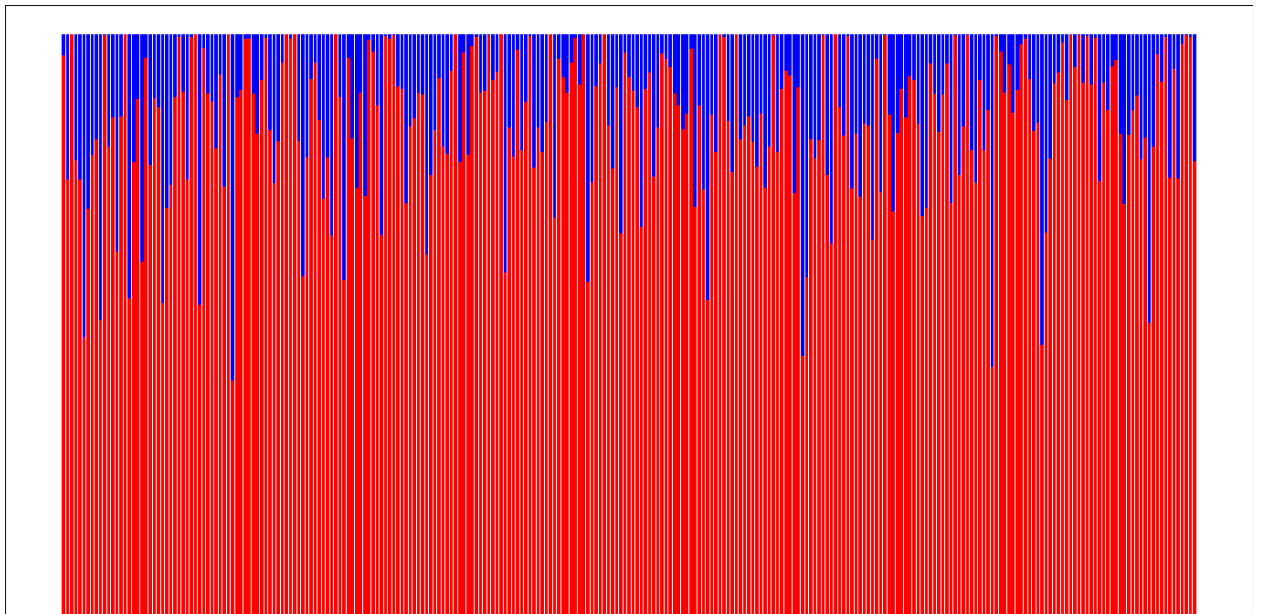


Рисунок 5.14 – Точність отримана на рівні об'єктів

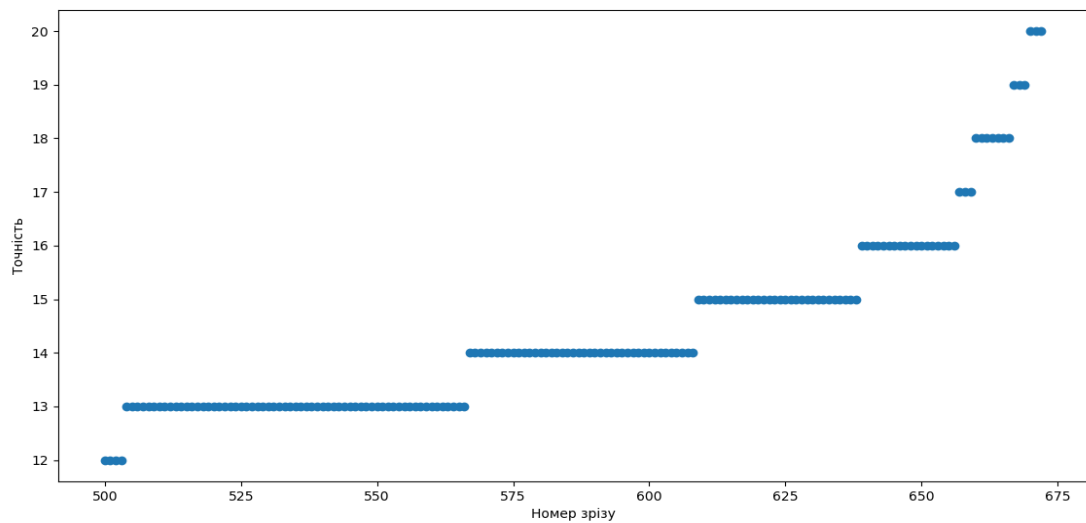


Рисунок 5.15 – Кількість зрізів в об'єктах

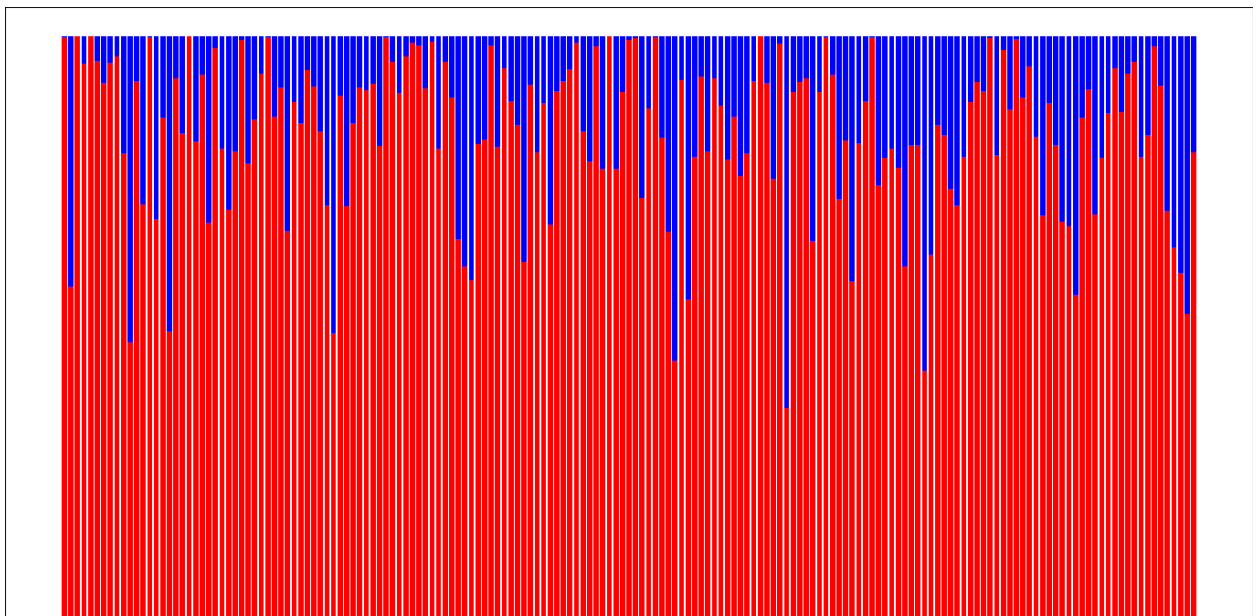


Рисунок 5.16 – Точність отримана на рівні об'єктів

Експеримент 8. Розподіл між неправильними відповідями.

Розроблена система пристосована для класифікації трьох класів об'єктів, як було показано вище - система інколи робить помилки, і відсоток цих помилок був визначений в попередніх експериментах. Але в них не було наведено інформації про те, в якому відсотковому співвідношенні розділяються неправильні відповіді між класами. Тобто, якщо при визначенні класу Healthy Controls система в 20%

випадків надала неправильний результат, то скільки з цих 20% належить до класу Alzheimer, а скільки до класу MCI.

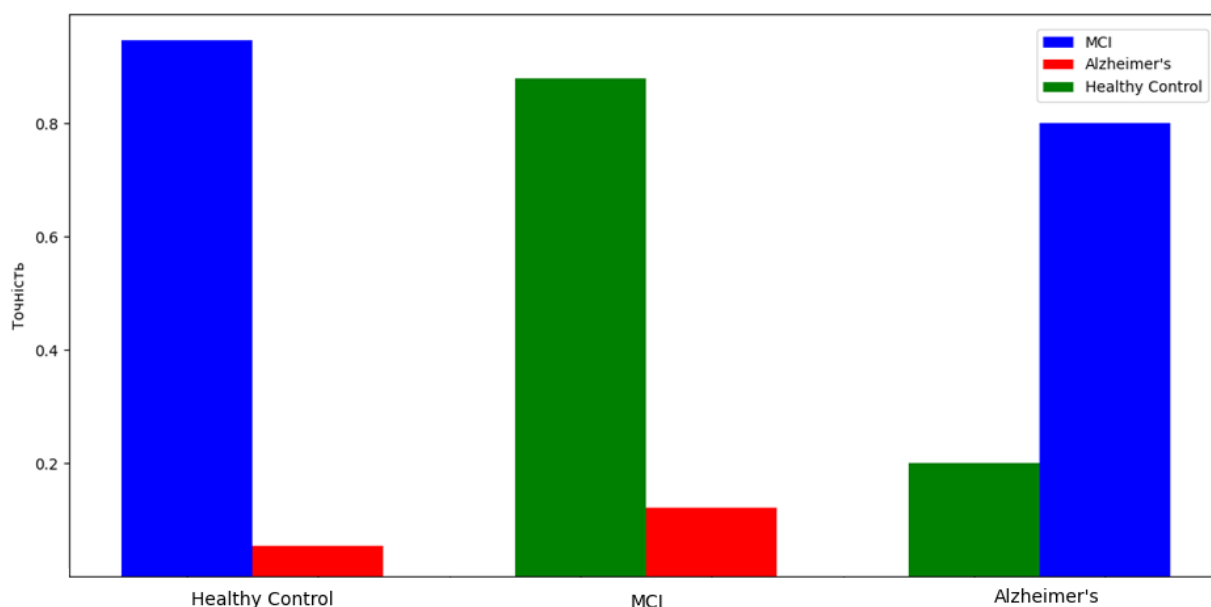


Рисунок 3.17 – Розподіл відповідей між класами, коли система надала неправильну відповідь (зліва на право розподіл коли система мала надати відповідь про приналежність об'єкту до 1-го, 2-го та 3-тього класу, відповідно)

Експеримент 9. Оцінка коректності тестових результатів.

Враховуючи те, що при формуванні тренувальної та тестової вибірки зрізи для кожної з них вибиралися із загального масиву випадковим чином, це призвело до ситуації, коли сусідні зрізи, до зрізу, який опинився в тестовій вибірці, могли опинитися в тренувальній вибірці. Враховуючи невелику відмінність між сусідніми зрізами, а також достатньо велику відмінність між тренувальними та тестовими результатами виникла необхідність в оцінці коректності тестових результатів. Для цього були відфільтровані із тестової вибірки, окремо, зрізи до яких прилягає один і два зрізи із тренувального набору, відповідно.

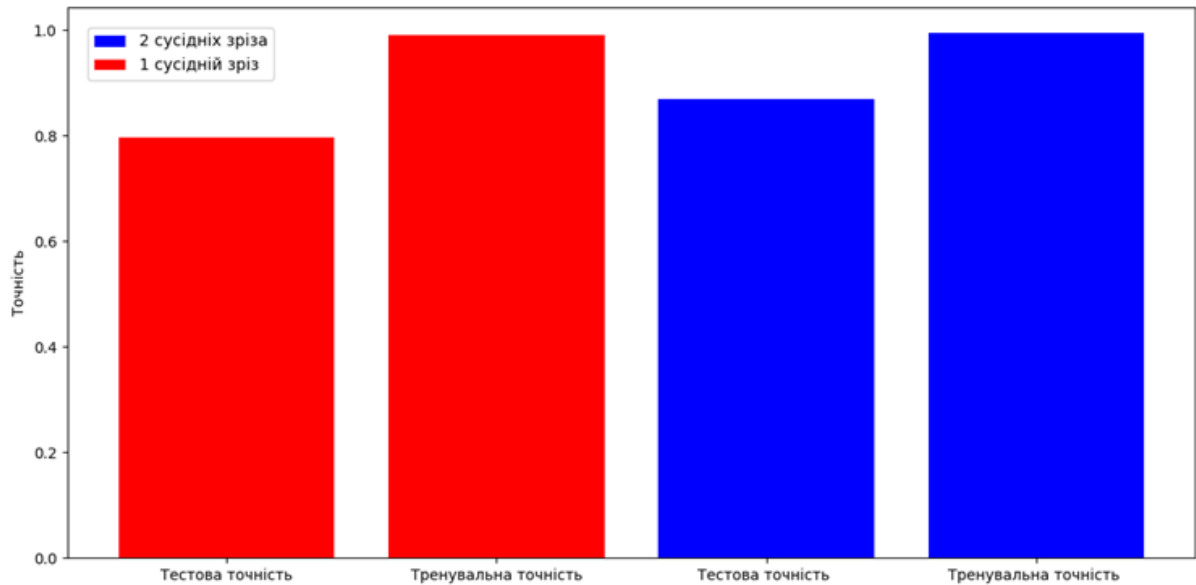


Рисунок 5.18 – Оцінка коректності тестових результатів

5.3 Висновки до п'ятого розділу

- 1) Зважаючи на точність тренувальній вибірці, яка склала майже 100% і точність в тестовій вибірці можна стверджувати, що система, особливо під час останніх тренувальних партій почала перенавчатися, що не дозволило їй збільшити точність на тестовій вибірці.
- 2) Із проведених експериментів очевидно, що результати отримані за допомогою розробленого класифікатора можуть бути тільки вхідними даними для ансамблю прийняття рішення.
- 3) Формувати тренувальну та тестову вибірку потрібно на рівні унікальних об'єктів (тобто на рівні 3D зображень мозку). Формування вибірок на рівні зрізів, в випадку перенавчання системи, може призводити до відгуку цього ефекту в тестовій вибірці.

- 4) Виходячи із експерименту 5, 6 аналізувати для аналізу в реальній медичній практиці потрібно використовувати зрізи з 40 по 60.

ВИСНОВКИ

Сучасні методи класифікації зображень на основі штучних нейронних мереж дозволяють побудувати систему автоматичної діагностики хвороби Альцгеймера на основі знімків МРТ, одним з підтверджень чого стали результати проведеного в цій роботі дослідження. Втім, як було показано, нейромережева модель може лише виступати одним з компонентів цієї системи, навколо якого повинен бути побудований ансамбль для прийняття рішень. Тобто системи, яка, з огляду на специфічність розв'язуваної задачі, дозволить максимально точно інтерпретувати результати, що формуються нейронною мережею.

Досить серозною перешкодою для поліпшення коректності роботи нейромережевої моделі в задачі діагностики хвороби Альцгеймера є порівняно невелика база даних знімків МРТ.

Враховуючи результати отримані в даній роботі – автоматичні системи діагностики мають стати невід'ємною частиною в сучасній медичній практиці.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Pouryamout, Laura, et al. "Economic Evaluation of Treatment Options in Patients with Alzheimer's Disease." *Drugs* 72.6 (2012): 789-802.
2. Асоціація Альцгеймеру. – Режим доступу: <https://www.alz.org/facts/>
3. Всесвітня організація охорони здоров'я. Департамент по економічним та соціальним питанням. – Режим доступу: <https://population.un.org/wpp/>
4. Ferri, Cleusa P., et al. "Global prevalence of dementia: a Delphi consensus study." *The lancet* 366.9503 (2005): 2112-2117.
5. Arizaga, R. L. "Dementia: a multidisciplinary approach." (2005).
6. Fratiglioni, L., et al. "Handbook of Neuropsychology." (2001).
7. Brookmeyer, Ron, et al. "Survival following a diagnosis of Alzheimer disease." *Archives of neurology* 59.11 (2002): 1764-1767.
8. Meek, Patrick D., E. Kristin McKeithan, and Glen T. Schumock. "Economic considerations in Alzheimer's disease." *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 18.2P2 (1998): 68-73.
9. Wimo, Anders, Bengt Winblad, and Linus Jönsson. "An estimate of the total worldwide societal costs of dementia in 2005." *Alzheimer's & Dementia* 3.2 (2007): 81-91.
10. Schoenmakers, Birgitte, et al. "The burden of family care giving in dementia." *Archives of public health* 62.3-4 (2004): 157-172.
11. Casey, David A., Demetra Antimisiaris, and James O'Brien. "Drugs for Alzheimer's disease: are they effective?." *Pharmacy and Therapeutics* 35.4 (2010): 208.
12. Ballard, Clive et al. "Alzheimer's disease." *Lancet (London, England)* vol. 377,9770 (2011): 1019-31. doi:10.1016/S0140-6736(10)61349-9
13. Tiraboschi, P et al. "The importance of neuritic plaques and tangles to the development and evolution of AD." *Neurology* vol. 62,11 (2004): 1984-9. doi:10.1212/01.wnl.0000129697.01779.0a

14. Atri, Alireza, et al. "Long-term course and effectiveness of combination therapy in Alzheimer's disease." *Alzheimer disease and associated disorders* 22.3 (2008): 209.
15. Стадії хвороби Альцгеймера: як прогресує хвороба. – Режим доступу до ресурсу: <https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/in-depth/alzheimers-stages/art-20048448>
16. Billones, Ciprian D., et al. "DemNet: A convolutional neural network for the detection of Alzheimer's disease and mild cognitive impairment." *2016 IEEE Region 10 Conference (TENCON)*. IEEE, 2016.
17. Платформа ImageNet. – Режим доступу: <http://www.image-net.org/>
18. Найкращі робочі місця в США. – Режим доступу: <https://www.indeed.com/lead/best-jobs-2019>
19. Звіт по світовому ринку машинного навчання. – Режим доступу: <https://www.marketresearchfuture.com/reports/machine-learning-market-2494>
20. Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." *Neural networks* 61 (2015): 85-117.
21. McCulloch, Warren S., and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity." *The bulletin of mathematical biophysics* 5.4 (1943): 115-133.
22. Насонов Д. Н., Местная реакция, закон «все или ничего» и автоматическая деятельность, «Изв. АН СССР. Серия биологическая», 1948, № 4, с. 381—92.
23. Потенціал дії. – Режим доступу: https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D1%82%D0%B5%D0%BD%D1%86%D0%B8%D0%B0%D0%BB_%D0%B4%D0%B5%D0%B9%D1%81%D1%82%D0%B2%D0%B8%D1%8F
24. Hebb, Donald Olding. *The organization of behavior: a neuropsychological theory*. J. Wiley; Chapman & Hall, 1949.
25. Farley, B. W. A. C., and W. Clark. "Simulation of self-organizing systems by digital computer." *Transactions of the IRE Professional Group on Information Theory* 4.4 (1954): 76-84.

26. Rosenblatt, Frank. "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological review* 65.6 (1958): 386.
27. Галушкин, Александр Иванович. *Синтез многослойных систем распознавания образов*. Энергия, 1974.
28. Werbos, P. J. "Beyond regression: New tools for prediction and analysis in the behavioral sciences. Ph. D. thesis, Harvard University, Cambridge, MA, 1974." (1974).
29. McClelland, James L., David E. Rumelhart, and PDP Research Group. "Parallel distributed processing." *Explorations in the Microstructure of Cognition* 2 (1986): 216-271.
30. Барцев, С. И., and В. А. Охонин. "Адаптивные сети обработки информации." *Красноярск: Ин-т физики СО АН СССР* 59Б (1986).
31. LeCun, Yann, et al. "Backpropagation applied to handwritten zip code recognition." *Neural computation* 1.4 (1989): 541-551.
32. Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).
33. Do, Quang, and G. Grudnitski. "A neural network analysis of the effect of age on housing values." *Journal of Real Estate Research* 8.2 (1993): 253-264.
34. Hopfield, John J. "Neural networks and physical systems with emergent collective computational abilities." *Proceedings of the national academy of sciences* 79.8 (1982): 2554-2558.
35. Wen, Junhao, et al. "Convolutional Neural Networks for Classification of Alzheimer's Disease: Overview and Reproducible Evaluation." *Medical Image Analysis* (2020): 101694.
36. Farooq, Ammarah, et al. "A deep CNN based multi-class classification of Alzheimer's disease using MRI." *2017 IEEE International Conference on Imaging systems and techniques (IST)*. IEEE, 2017.
37. Gunawardena, K. A. N. N. P., R. N. Rajapakse, and N. D. Kodikara. "Applying convolutional neural networks for pre-detection of alzheimer's disease from

- structural MRI data." *2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*. IEEE, 2017.
- 38.Hon, Marcia, and Naimul Mefraz Khan. "Towards Alzheimer's disease classification through transfer learning." *2017 IEEE International conference on bioinformatics and biomedicine (BIBM)*. IEEE, 2017.
 - 39.Islam, Jyoti, and Yanqing Zhang. "Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks." *Brain informatics* 5.2 (2018): 2.
 - 40.Qiu, Chengxuan, and Laura Fratiglioni. "Aging without dementia is achievable: current evidence from epidemiological research." *Journal of Alzheimer's Disease* 62.3 (2018): 933-942.
 - 41.Wang, Shui-Hua, et al. "Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling." *Journal of medical systems* 42.5 (2018): 85.
 - 42.He, Kaiming, et al. "Identity mappings in deep residual networks." *European conference on computer vision*. Springer, Cham, 2016.
 - 43.Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
 - 44.Liu, Manhua, et al. "Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis." *Neuroinformatics* 16.3-4 (2018): 295-308.
 - 45.Li, Fan et al. "Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks." *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society* vol. 70 (2018): 101-110. doi:10.1016/j.compmedimag.2018.09.009
 - 46.Lian, Chunfeng, et al. "Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI." *IEEE transactions on pattern analysis and machine intelligence* (2018).
 - 47.Liu, Mingxia, et al. "Landmark-based deep multi-instance learning for brain disease diagnosis." *Medical image analysis* 43 (2018): 157-168.

48. Dickerson, Bradford C., et al. "MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease." *Neurobiology of aging* 22.5 (2001): 747-754.
49. Salvatore, Christian, et al. "Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach." *Frontiers in neuroscience* 9 (2015): 307.
50. Bäckström, Karl, et al. "An efficient 3D deep convolutional network for Alzheimer's disease diagnosis using MR images." *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018.
51. Basaia, Silvia et al. "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks." *NeuroImage. Clinical* vol. 21 (2019): 101645. doi:10.1016/j.nicl.2018.101645
52. Hosseini-Asl, Ehsan et al. "Alzheimer's disease diagnostics by a 3D deeply supervised adaptable convolutional network." *Frontiers in bioscience (Landmark edition)* vol. 23 584-596. 1 Jan. 2018, doi:10.2741/4606
53. Billones, Ciprian D., et al. "DemNet: A convolutional neural network for the detection of Alzheimer's disease and mild cognitive impairment." *2016 IEEE Region 10 Conference (TENCON)*. IEEE, 2016.
54. Wen, Junhao, et al. "Convolutional Neural Networks for Classification of Alzheimer's Disease: Overview and Reproducible Evaluation." *Medical Image Analysis* (2020): 101694.
55. Бібліотека TensorFlow. – Режим доступу: <https://www.tensorflow.org/>
56. Платформа ADNI. – Режим доступу: <http://adni.loni.usc.edu/>
57. Бібліотека OpenCV. – Режим доступу: <https://opencv.org/>